Geographical Roots of the Coevolution of Cultural and Linguistic Traits

Oded Galor,^{1*} Ömer Özak,² Assaf Sarid³

¹Department of Economics, Brown University; NBER, CEPR, IZA, CES-Ifo

²Department of Economics, Southern Methodist University; IZA

³Department of Economics, University of Haifa

*To whom correspondence should be addressed; E-mail: Oded.Galor@brown.edu.

August 3, 2020

This research explores the geographical origins of the coevolution of cultural and linguistic traits. It establishes the existence of common roots for: (i) the presence of sex-based grammatical gender and the prevalence of gender bias, (ii) the existence of politeness distinctions and power distance, and (iii) the structure of the future tense and long-term orientation. The study documents that geographical characteristics that generated gender gaps in agricultural productivity, and led to gender bias in society, fostered the existence of sex-based grammatical gender. Further, ecological diversity, and its effect on the formation of hierarchical societies and the rise in power distance, were conducive to the emergence of politeness distinctions. Finally, higher caloric crop return, and its impact on long-term orientation, contributed to cross-language variations in the structure of the future tense.

One-Sentence Summary: This research identifies common geographical roots for the co-evolution of various complementary cultural and linguistic traits.

Short Title: Roots of Cultural and Linguistic Traits

Evidence suggests that regional variations in the geographical environment in the distant past have contributed to the differential formation of cultural traits and their lasting effect on comparative economic development across countries, regions and ethnic groups (1–4). In light of the conceivable coevolution of cultural and linguistic characteristics in the course of human history, emerging linguistic traits may have reinforced the persistent effect of cultural factors on the process of development. Thus, the exploration of these joint evolutionary processes, and their potential common geographical roots, can be instrumental for the understanding of the long shadow of history on comparative economic development and the vast inequality across of nations.

This research contributes to one of the main controversies in cultural evolution and linguistics. Unlike the hypothesis that observed correlations between cultural and linguistic traits reflect random associations, this research advances the hypothesis and documents empirically that geographical characteristics which were conducive to the emergence of fundamental cultural traits, have contributed to the evolution of complementary linguistic traits that have reinforced the impact of these cultural characteristics on the process of development.

This novel hypothesis rests upon several fundamental building blocks. First, in light of the pivotal role of language in the diffusion of knowledge and the transmission of values, emerging linguistic characteristics have conceivably affected the diffusion of cultural values, reinforcing existing cultural traits and their intergenerational transmission (5). Second, the forces of natural selection across language structures have plausibly generated an evolutionary advantage to linguistic traits that reflected and reinforced the dominating cultural traits (6). Third, regional differences in geographical characteristics that have contributed to the emergence of variations in cultural traits have conceivably contributed to the evolution of cross-language variations in complementary linguistic traits.

The research explores this general hypothesis focusing on three cultural traits that are fundamental to the process of development, namely (i) the prevalence of gender bias in society (e.g., beliefs about gender roles, gender gaps in labor force participation, occupation, and education), (ii) the pervasiveness of hierarchical orientation (e.g., beliefs about the role of authority and government, respect for authority and parents, power distance, autocracy), and (iii) the degree of long-term orientation in society (e.g., the ability to delay gratification, patience, human capital formation, savings) (2, 7). The analysis explores these cultural traits in relation to complementary linguistic traits have been proposed by socio-linguists (5, 8–14). Specifically, it explores the coevolution of (i) the presence of sex-based grammatical gender and the prevalence

of gender bias in society, (ii) the existence of politeness distinctions and the pervasiveness of hierarchical orientation, and (iii) the existence of periphrastic future tense and degree of long-term orientation in society.

Geography and the Coevolution of Cultural & Linguistic Traits

In line with the proposed hypothesis, these cultural traits are associated with their hypothesized complementary linguistic traits. In particular, as depicted in Figure 1 and further analyzed in Supplementary Materials S2, S4.D, S5.D, and S6.D, there exists a highly significant positive association, across language groups, between each cultural trait and its linguistic complement.

Figure 1: Contemporary Cultural & Linguistic Traits

The figure depicts the association between linguistic and cultural traits across languages as derived in Table S2.1. Panel (A) depicts the increase in the prevalence of gender bias in societies where sex-based grammatical gender is present relative to societies in which it is absent. Panel (B) depicts the increase in the prevalence of hierarchical structure in societies where politeness distinction is present relative to societies in which it is absent. Panel (C) depicts the increase in the prevalence of long-term orientation in societies where periphrastic future tense is present relative to societies in which it is absent. See Supplemental Material S1 for the definition and sources of the variables used in this figure. Each colored rectangle represents the impact (measured in standard deviations) of the existence of a linguistic trait on the prevalence of a cultural trait, accounting for geographical and climatic factors, as well as regional fixed-effects (Supplemental Material S4.D, S5.D, and S6.D).

In view of the contribution of geographical characteristics to the evolution of cultural traits, this research advances the hypothesis that common geographical roots have governed the coevolution of complementary cultural and linguistic traits. In particular, in societies characterized by a sexual division of labor and consequently by the existence of gender biases, linguistic traits that had fortified the existing gender biases had plausibly emerged and persisted over time. Indeed, socio-linguists as well as other scholars have argued that gender bias can be reinforced by a sex-based grammatical gender system (11–13). Thus, geographical characteristics conducive

to the adoption of the plow, which required physical upper-body strength to pull it, and therefore generated a gender gap in productivity fostered both the emergence of distinct gender roles in society (3, 15) and the emergence and the presence of sex-based grammatical gender in the course of human history.

Similarly, the proposed hypothesis suggests that in societies characterized by hierarchical orientation (e.g., obedience, conformity, and power distance), linguistic traits that had reinforced existing hierarchical structures and cultural norms had emerged and persisted. In particular, politeness distinctions in pronouns (e.g., the differential use of “tu” and “usted” in the Spanish language, “Du” and “Sie” in German, and “tu” and “vous” in French) had plausibly appeared and endured in hierarchical societies in order to mitigate the coordination costs in the interaction between individuals from various social strata (5, 16, 17). Thus, geographical characteristics, such as ecological diversity that had been conducive to the emergence of trade and economic specialization (18), contributed to the emergence and coevolution of hierarchical societies and politeness distinctions.

Finally, the research suggests that in societies characterized by long-term orientation (i.e., the ability to delay gratification), a structure of the future tense that reinforced long-term oriented behavior may have emerged and persisted over time. Indeed, languages characterized by a *periphrastic* rather than *inflectional* future tense (19), a form that expresses a desire, an intention, an obligation, a commitment as well as a movement towards a goal (9), had plausibly appeared and endured in long-term oriented societies. Thus, geographical characteristics that were conducive to higher caloric crop return, which triggered the emergence and the prevalence of long-term orientation (4), were conducive to the existence of a long-term oriented structure of the future tense.

Data

The analysis is based on data on individual languages, their linguistic traits, the cultural and ethnographic characteristics of their speakers, as well as the geographical characteristics of their homelands. The dataset includes information on (i) the existence of sex-based grammatical gender in 217 contemporary languages (74 language families), (ii) the existence of politeness distinctions in pronouns in 207 contemporary languages (69 language families), and (iii) the existence of periphrastic future tense for 275 contemporary languages (76 language families). The global distribution of these traits is depicted in Figure S1.1.

The proposed geographical roots are measured in the contemporary homeland of each language (i.e., where it was spoken before 1500CE) and the ancestral homeland of its proto-language (i.e., *Urheimat*) (Supplemental Material S1). First, the potential caloric crop return (4) measures the potential daily calories that could be produced by cultivating the crops that would maximize caloric yields given crop growth cycles in the pre-1500CE era (i.e., prior to the Columbian Exchange). Second, the potential historical suitability of land for the usage of the plow is measured by the potential caloric yield of crops that are unsuitable for the use of the plow in the pre-1500CE period (20). Third, the extent of ecological diversity is measured by a Herfindahl index of the share of each territory that is located in different ecological zones (18).

As elaborated in Supplemental Material S1, each language is linked to: (i) its linguistic traits, as reported in *The World Atlas of Language Structures* (WALS) (21), (ii) the geographical characteristics of its contemporary homeland, as mapped in *Ethnologue* (22), (iii) the geographical characteristics of the hypothesized ancestral homeland of its proto-language (i.e., its *Urheimat*), as determined following the method of (23), (iv) the ethnographic characteristics of the groups that speak the language, based on the *Ethnographic Atlas* (24), and (v) the cultural characteristics of its contemporary speakers, based on the *World Values Survey* (25) as well as other sources listed in Supplementary Material S1.

Results

This research documents the common geographical roots of the coevolution of these cultural and linguistic traits, as reported in Table 1, Figure 2 and Supplemental Material S4-S6. First, the analysis suggests that suitability of land for the use of the plow is associated with the emergence of sex-based grammatical gender (Figure 2-A1) and the prevalence of gender bias in society (Figure 2-A2). Second, it documents that ecological diversity is associated with the emergence of politeness distinctions (Figure 2-B1) and the presence of hierarchical societies (Figure 2-B2). Third, it demonstrates that caloric crop return is associated with the emergence of the periphrastic future tense (Figure 2-C1) and the pervasiveness of long-term orientation (Figure 2-C2).

In particular, as reported in Table 1, the analysis establishes that after accounting for the confounding effects of climatic and geographical characteristics, as well as regional time-invariant characteristics: (i) a one standard deviation increase in the caloric suitability for the plow is associated with a 23 percentage point increase in the probability that a language has sex-based

Figure 2: Common Geographical Roots of Cultural & Linguistic Traits

The figure depicts the common geographical roots of cultural and linguistic traits. The left panels (A) depict the contribution of the suitability of the land for the use of the plow to the presence of sex-based grammatical gender in a language (A1) and gender bias in its speech community (A2). The middle panels (B) depict the contribution of ecological diversity to the presence of politeness distinction in a language (B1) and hierarchical structure in its speech community (B2). The right panels (C) depict the contribution of caloric crop return to the the presence of periphrastic future tense in a language (C1) and long-term orientation in its speech community (C2). The scatter plots represent the partial correlation between each linguistic or cultural trait and its geographical root after accounting for a set of basic confounders in Tables S5.4, S6.4, S4.5.

grammatical gender (Column 3). (ii) a one standard deviation increase in ecological diversity is associated with a 10 percentage point increase in the probability that a language has politeness distinctions (Column 6). (iii) a one standard deviation increase in the caloric crop return is associated with a 11 percentage point increase in the probability that a language has a periphrastic future tense (Column 9).

Importantly, the associations in columns 3 and 9 cannot reflect reverse causality from linguistic traits, and their associated cultural traits, to these geographical attributes, since the analysis exploits variations in potential crop return (associated with agro-climatic conditions that are orthogonal to human intervention), rather than the potentially endogenous, actual crop return,

and the potential suitability for the usage of the plow, rather than its actual usage. Furthermore, the analysis addresses many identification issues encountered in this type of linguistic studies (26) (see Supplemental Materials S7, S8, S9). Although additional research will be needed to further strengthen the proposed hypothesis, the main results are quite robust to using alternative statistical models (probit, logistic, multilevel mixed-effects logistic), accounting for spatial and historical dependence within language families and genres (multilevel mixed-effects logistic, clustering of standard errors and spatial-autocorrelation), and corrections to potential biases due to selection on observables or omitted variables.

Table 1: Geographic Origins of Language Structures (OLS)

	Existence of Language Structure								
	Gender			Politeness			Future		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Caloric Plow Suitability	0.13** (0.06)	0.29*** (0.06)	0.23*** (0.08)						
Ecological Diversity				0.13*** (0.03)	0.10** (0.04)	0.10** (0.04)			
Caloric Crop Return							0.06** (0.03)	0.09*** (0.03)	0.11*** (0.03)
Average Caloric Yield	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Geographical Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Precipitation Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Temperature Controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Unproductive Period	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Regional FE	No	No	Yes	No	No	Yes	No	No	Yes
Adjusted- R^2	0.03	0.15	0.21	0.14	0.21	0.32	0.01	0.07	0.11
Observations	217	217	217	198	198	198	275	275	275

Notes: This table documents the significant positive association between the geographical and linguistic traits. The coefficients are generated by a linear probability model, which accounts for geographic and climatic characteristics and regional fixed-effects. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients are comparable and measure the impact of a one standard deviation increase in the independent variable on the probability that a language has a specific linguistic trait. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes p-value ($p < 0.01$), ** p-value ($p < 0.05$), and * p-value ($p < 0.1$), all for two-sided hypothesis tests.

Nevertheless, the observed associations between geographical and linguistic traits may reflect the sorting of individuals into geographical niches that complement their linguistic and cultural traits. While sorting would not affect the nature of the association between these geographical and linguistic characteristics (i.e., variations in geographical characteristics across the

globe would still be the origin of cross-language variation in linguistic traits), it would weaken the underlying mechanism. Hence, to alleviate concerns about the potential sorting of individuals into regions that complement their linguistic traits, as well as to explore the historical depth of the association between these geographical characteristics and language structures, the analysis exploits the well-documented observation that daughter languages within a language family originated from a common proto-language (27–29).

Figure 3: Geographical Roots and Other Language Structures (WALS)

The figure explores the orthogonality of the 3 geographical characteristics to other linguistic traits reported in WALS. The left panels (A) depict the significant contribution of the suitability of the land for the use of the plow for 3 forms of grammatical gender (in blue) and its orthogonality to most language structures (A1: 1-68 and A2: 69-132). The middle panels (B) depict the significant contribution of ecological diversity for the presence of politeness distinctions (in yellow) and its orthogonality to most language structures (B1: 1-68 and B2: 69-132). The right panels (C) depict the significant contribution of caloric crop return for the presence of periphrastic future tense (in red) and its orthogonality to most language structures (C1: 1-68 and C2: 69-132). The bullets depict the estimated effect with its 95% confidence interval around it (see Supplemental Material S4-S6 for further details, especially Tables S14.1, S14.3, S14.7 and Figures S14.3, S14.4, S14.5).

Table 2 documents that geographical characteristics in the Urheimat of a language (i.e., the homeland of its proto-language), rather than its contemporary homeland govern the association with sex-based grammatical gender and periphrastic future tense. Thus, these associations sug-

Table 2: Persistent Effect of Urheimat Characteristics
Languages Outside Urheimat

	Existence of Language Structure					
	Migratory Distance to Urheimat					
	Any Distance			At Least 1 Week		
	Gender	Politeness	Future	Gender	Politeness	Future
	(1)	(2)	(3)	(4)	(5)	(6)
Homeland's Caloric Plow Suitability	0.11 (0.07)			0.13** (0.05)		
Urheimat's Caloric Plow Suitability	0.31*** (0.10)			0.47** (0.21)		
Homeland's Ecological Diversity		0.07* (0.04)			0.10*** (0.02)	
Urheimat's Ecological Diversity		0.05 (0.03)			0.07 (0.05)	
Homeland's Caloric Crop Return			0.01 (0.05)			-0.01 (0.06)
Urheimat's Caloric Crop Return			0.17** (0.08)			0.24** (0.09)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes
Homeland Geographical Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Urheimat Geographical Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted- R^2	0.41	0.41	0.17	0.50	0.49	0.15
Observations	195	180	255	131	126	195
Language Families	57	54	59	29	28	32

Notes: This table documents the importance of geographical characteristics in the Urheimat, rather than in the contemporary homeland, to the presence of linguistic traits in daughter languages located outside of the Urheimat. The coefficients are generated by a linear probability model, which accounts for geographic and climatic characteristics and regional fixed-effects. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length. Columns (1)-(3) consider daughter languages which are located anywhere outside of the Urheimat, while column (4)-(6) those which are at least a week of migratory distance from the Urheimat. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients are comparable and measure the impact of a one standard deviation in the independent variable on the probability that a language has a specific linguistic trait. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes p-value ($p < 0.01$), ** p-value ($p < 0.05$), and * p-value ($p < 0.1$), all for two-sided hypothesis tests.

gest that these linguistics traits emerged in the Urheimat prior to the demic diffusion of daughter languages and therefore cannot reflect sorting but rather the coevolution of cultural and linguistic traits prior to this migration process. These results are in line with linguistic evidence on the temporal stability of sex-based grammatical gender and the future tense (30–32). In contrast,

consistent with evidence about the greater adaptability of politeness distinctions to environmental and political changes (30–32), the results suggest that politeness distinctions may have emerged primarily during the demic diffusion of the language and sorting cannot be excluded. (SM S11.A explores robustness to location of Urheimat).

The associations between geographical and linguistic traits may reflect a broader association between each of these geographical factors and a wide range of linguistic traits. However as depicted in Figure 3, alternative language structures (as reported by WALS), such as the existence of other temporal structures (e.g., the past tense and the perfect tense), as well as non-temporal structures (e.g., possessive, evidentiality, and the number of consonants and colors), are generally not associated with each of the proposed geographical roots. As further discussed in Supplementary Material S4-S6, some of the exceptions lend additional credence to the proposed hypothesis, since they are associated with the same cultural characteristics.

Mechanism

The analysis depicted in Figure 4 further examines the potential mediating channels that may have governed the coevolution of these cultural and linguistic traits (See also Supplemental Material S4-S6, Tables S4.3, S5.3, S6.3). First, the findings suggests that the aboriginal use of the plow across ethnic groups is the mediating channel through which caloric plow suitability may have affected the presence of sex-based grammatical gender. Second, the results indicate that the level of jurisdictional hierarchy which captures the hierarchical nature of these ethnic groups may be a mediating channel through which ecological diversity has affected the presence of politeness distinctions. Third, the results suggest that the practice of agriculture by ethnic groups appears to be the mediating channel through which the caloric crop return has affected the presence of periphrastic future tense.

Discussion

This research establishes the geographical determinants of the coevolution of cultural and linguistic traits in the course of human history. It contributes to the growing literature on the evolution of languages (33–35), providing (in contrast to received wisdom (26)) evidence that the structure of languages is geographically based. The study presents evidence for the complementarity between linguistic and cultural traits, lending credence to the viewpoint that language and culture have co-evolved (36). Thus, it provides geographical foundations for the link be-

Figure 4: Geographical Roots of Linguistic Traits: Mechanisms

The figure depicts the hypothesized mechanism through which geographical characteristics affect linguistic traits. Panel (A1) depicts the hypothesized link between caloric plow suitability, the usage of the plow and the emergence of sex-based grammatical gender. Panel (A2) depicts the probability of the existence of sex-based grammatical gender between ethnic groups that use the plow and those which do not. Panel (B1) depicts the hypothesized link between ecological diversity, the level of hierarchy and the emergence of politeness distinctions. Panel (B2) depicts the probability of the existence of politeness distinctions across ethnic groups with different levels of hierarchy. Panel (C1) depicts the hypothesized link between caloric crop return, the presence of agriculture and the emergence of periphrastic future tense. Panel (C2) depicts the probability of the existence of periphrastic future tense across agricultural and non-agricultural ethnic groups. See also Supplemental Material Table S13.1 and sections S4-S6.

tween linguistic traits and socio-economic behavior (5, 6, 12, 14, 37–40). Moreover, it sheds a new light on the stability of some language structures (30–32, 41, 42), underlying the importance of the geographical characteristics at the Urheimat in observed contemporary structures. Finally, the analysis sheds additional light on the geographical roots of comparative development (43, 44), and on the interaction between the evolution of human traits and the process of development (45).

References

1. L. Guiso, P. Sapienza, L. Zingales, *American Economic Review* **94**, 526 (2004).

2. L. Guiso, P. Sapienza, L. Zingales, *Journal of Economic Perspectives* **20**, 23 (2006).
3. A. Alesina, P. Giuliano, N. Nunn, *The Quarterly Journal of Economics* **128**, 469 (2013).
4. O. Galor, Ö. Özak, *American Economic Review* **106**, 3064 (2016).
5. P. Brown, S. C. Levinson, *Politeness: Some universals in language usage*, vol. 4 (Cambridge university press, 1987).
6. S. Roberts, J. Winters, *Psychology of Language and Communication* **16**, 89 (2012).
7. G. H. Hofstede, G. J. Hofstede, M. Minkov, *Cultures and organizations: software of the mind : intercultural cooperation and its importance for survival* (McGraw-Hill, New York, 2010), third edn.
8. J. L. Bybee, W. Pagliuca, *Papers from the 7th international conference on historical linguistics* (Amsterdam: John Benjamins, 1987), pp. 109–122.
9. J. L. Bybee, Ö. Dahl, *Studies in language* **13**, 51 (1989).
10. J. L. Bybee, R. Perkins, W. Pagliuca, *The evolution of grammar* (1994).
11. R. Lakoff, *Language in society* **2**, 45 (1973).
12. P. Eckert, S. McConnell-Ginet, *Language and gender* (Cambridge University Press, Cambridge, 2003).
13. D. Stahlberg, F. Braun, L. Irmen, S. Sczesny, *Social communication* pp. 163–187 (2007).
14. R. Brown, A. Gilman, *Style in Language*, T. A. Sebeok, ed. (MIT Press, 1960), chap. 12.
15. E. Boserup, *Woman's role in economic development* (St. Martin's Press, New York, 1970).
16. R. Brown, A. Gilman, *Language in society* **18**, 159 (1989).
17. J. Helmbrecht, *Pragmatics and beyond New Series* pp. 185–202 (2003).
18. J. Fenske, *Journal of the European Economic Association* **12**, 612 (2014).
19. Ö. Dahl, V. Velupillai, *The World Atlas of Language Structures Online*, M. S. Dryer, M. Haspelmath, eds. (Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013).

20. F. L. Pryor, *Comparative Studies in Society and history* **27**, 727 (1985).
21. M. Dryer, Matthew S & Haspelmath, ed., *The World Atlas of of Language Structures Online*. (Leipzig: Max Planck Institute for Evolutionary Anthropology., 2013).
22. M. P. Lewis, G. F. Simons, C. D. Fennig, *Ethnologue: Languages of the world*, vol. 16 (SIL international Dallas, TX, 2009).
23. S. Wichmann, A. Müller, V. Velupillai, *Diachronica* **27**, 247 (2010).
24. G. P. Murdock, *Ethnology* pp. 109–236 (1967).
25. R. Inglehart, *et al.*, *World Values Survey: All Rounds - Country-Pooled Datafile* (JD Systems Institute, Madrid, 2014).
26. D. R. Ladd, S. G. Roberts, D. Dediu (2015).
27. R. Bouckaert, *et al.*, *Science* **337**, 957 (2012).
28. M. Pagel, Q. D. Atkinson, A. S. Calude, A. Meade, *Proceedings of the National Academy of Sciences* **110**, 8471 (2013).
29. S. Wichmann, *Diachronica* **34**, 79 (2017).
30. S. Wichmann, E. W. Holman, *Temporal stability of linguistic typological features* (Lincom Europa, 2009).
31. D. Dediu, S. C. Levinson, *PloS One* **7**, e45198 (2012).
32. S. J. Greenhill, *et al.*, *Proceedings of the National Academy of Sciences* **114**, E8822 (2017).
33. S. Pinker, P. Bloom, *Behavioral and brain sciences* **13**, 707 (1990).
34. M. H. Christiansen, S. Kirby, *Language evolution*, M. H. Christiansen, S. Kirby, eds. (Oxford University Press, 2003), vol. 3, pp. 1–15.
35. D. Bickerton, *Lingua* **117**, 510 (2007).
36. F. J. Newmeyer, *Studies In The Evolution Of Language* **3**, 58 (2003).

37. M. Hellinger, H. Bußmann, *Gender Across Languages*, M. Hellinger, H. Motschenbacher, eds. (John Benjamins Publishing, 2001), vol. 4, chap. 3, pp. 1–25.
38. M. K. Chen, *The American Economic Review* **103**, 690 (2013).
39. L. S. Davis, F. Abdurazokzoda, *Journal of Comparative Economics* **44**, 541 (2016).
40. V. Ginsburgh, S. Weber, *The Palgrave Handbook of Economics and Language*, Palgrave handbooks (Palgrave, 2018).
41. S. J. Greenhill, Q. D. Atkinson, A. Meade, R. D. Gray, *Proceedings of the Royal Society of London B: Biological Sciences* **277**, 2443 (2010).
42. D. Dediu, M. Cysouw, *PloS one* **8**, e55009 (2013).
43. J. M. Diamond, *Guns, germs, and steel: the fates of human societies* (W.W. Norton & Co., New York, 1997), first edn.
44. Q. Ashraf, O. Galor, *The American Economic Review* **103**, 1 (2013).
45. E. Spolaore, R. Wacziarg, *Handbook of Economic Growth* (Elsevier, 2013), vol. 2, p. 121.
46. R. Godoy, *et al.*, *Journal of Anthropological Research* pp. 179–202 (2004).
47. Y. Hernryadin, K. Kotani, T. Saijo, *Land Economics* **96**, 132 (2020).
48. T. A. Kohler, *et al.*, *Nature* **551**, 619 (2017).

Acknowledgments

The authors are grateful to Alberto Alesina, Joan Bybee, Justin Cook, Lewis Davis, Klaus Desmet, Delia Furtado, Paola Giuliano, Luigi Guiso, Stelios Michalopoulos, David Weil and Søren Wichmann, as well as conference participants at the NBER Culture and Institutions Meeting; “Annual Meetings of the American Economic Association”; “Annual Meetings of the Southern Economic Association”; “Language Skills for Economic and Social Inclusion Colloquium”, Humboldt-University Berlin; “Deep Rooted Factors in Comparative Development”; “The Long Shadow of History: The Role of Geographical, Institutional, Cultural, and Human Factors in Comparative Economic Development”, The Pinhas Sapir Economic Policy Forum; Zeuthen Lectures, Copenhagen; and seminar participants at Ben-Gurion, Brown, Clark, Haifa,

Southern Methodist, Tel-Aviv, and UC-Merced, for useful comments and discussions. Sarid acknowledges financial support from the Israel Science Foundation grant No. 1331/18.

Author contributions

O.G., Ö.Ö. and A.S. conceived the research idea, formulated the theory, analyzed the data and wrote the manuscript.

Competing interests

Authors declare no competing interests.

Supplementary materials

Methods & Materials

S1 Data

This section presents the data that is used in the empirical analysis of the geographical origins of language structures and the methodology that guided its construction. In particular, it introduces the data on cross-language variations in the structure of the future tense, the presence of sex-based grammatical gender, and the existence of politeness distinctions in pronouns, as well as the data on cross-regional variations in the hypothesized geographical determinants of these language structures.

The linguistic characteristics of each language are linked to: (i) the geographical characteristics of the contemporary homeland (i.e., the characteristics of the indigenous regions that spoke the language in the pre-colonial as well as in the post-colonial era),¹ (ii) the geographical characteristics of the ancestral homeland of the proto-language (i.e., the *Urheimat*), and (iii) the ethnographic characteristics of the groups that speak the language in the contemporary linguistic homeland. In particular, the linguistic data is mapped to the *Ethnologue* (22) in order to identify the geographical characteristics of the contemporary linguistic homeland, and is linked to the Ethnographic Atlas (24) and the Standard Cross Cultural Sample (51) in order to identify the ethnographic characteristics of the aboriginal speakers of each language. The mapping of these datasets is based on the ISO code of each language, which is available in all three datasets.

Furthermore, in order to explore the deep historical roots of language structures, each language is mapped to the ancestral homeland of its proto-language (*Urheimat*). In particular, for each language family, the paper identifies the location of the *Urheimat* based on a methodology developed by linguists (23). This method, as explained further in Appendix S11 follows a similar strategy employed by geneticists to determine the location of the origin of domestication of crops.²

¹This focus on the contemporary linguistic homeland is designed to capture the geographical roots of each daughter language (rather than the proto-language). For instance, Spain, rather than regions in Latin America and elsewhere that had adopted the Spanish language in the post-colonial era, is considered the contemporary linguistic homeland of the Spanish language.

²The analysis excludes language families for which there is no data on the location of the *Urheimat*, or if the language family is unclassified, or the language is a Creole or language isolate (i.e., which has no identified sister languages).

S1.A Dependent Variables: Language Structures

This subsection introduces the three language structures whose geographical origins will be explored: (i) periphrastic future tense, (ii) sex-based grammatical gender systems, and (iii) politeness distinctions in pronouns. In particular, based on The World Atlas of Language Structures (WALS) (21), it presents the data about the existence of each of these structures across languages.

The existence of a structure S in language ℓ , S_ℓ , is encoded such that:

$$S_\ell = \begin{cases} 1 & \text{if the structure exists in language } \ell \\ 0 & \text{if the structure does not exist in language } \ell. \end{cases}$$

S1.A.1 Periphrastic Future Tense

Languages differ in the structure of their future tense. In particular, linguists distinguish between languages that are characterized by an *inflectional* versus *periphrastic* future tense (19, 54, 55). Inflectional future tense is associated with verbs that display morphological variation (i.e., a change in the verb form that is associated with the future tense). In contrast, periphrastic future tense is characterized by roundabout or discursive phrases, such as ‘will’, ‘shall’, ‘want to’, ‘going to’ in the English language (8–10).

(9) argue that, unlike the inflectional future tense, the periphrastic future tense is formed by terms that express a desire, an intention, an obligation, a commitment as well as a movement towards a goal. In particular, in the English language, “*shall has developed from a main verb meaning ‘to owe’, will from a main verb meaning ‘to want’, and the source of be going to is still transparent*” (9, p.90). Moreover, “*intention and prediction are most commonly expressed by the periphrastic future, while the synthetic one is more common in generic statements, concessives, and suppositions*” (10, p.235). Inflectional futures “*also appear systematically (often obligatorily) in sentences which express clear predictions about the future (which are independent of human intentions and planning), whereas less grammaticalized constructions [i.e., periphrastic] often tend to be predominantly used in talk of plans and intentions – a fact which is explainable from the diachronic sources of future tenses*” (19, p.270). Thus, these views suggest that the use of periphrastic future tense conceivably reflects long-term oriented behavior.

The dataset on the existence of periphrastic future tense includes 275 contemporary lan-

Figure S1.1: The Global Distribution of Language Structures in Analysis

guages, originated from 76 language families.³ The classification of 222 of the languages in the sample is based on (19), as reported in WALS, and the remaining 53 languages are classified based on (54, 55). In particular, languages that have no inflectional future tense are coded as having periphrastic future, while those that have an inflectional future are coded as not having a

³Importantly, 90% of the languages in the Ethnologue belong to these 76 language families.

periphrastic future.⁴ Figure S1.2 and Table S10.1 describe the prevalence and the spatial distribution of the structure of future tense in the dataset, portraying wide variations within regions in the existence of periphrastic future tense. In particular, in most regions about 50% of the languages in the sample are characterized by periphrastic future tense.

Figure S1.2: The Global Distribution of Periphrastic Future Tense

S1.A.2 Sex-Based Grammatical Gender Systems

Languages differ in the existence and the form of grammatical gender. In particular, languages that are characterized by sex-based grammatical gender classify nouns according to biological gender. The presence of sex-based grammatical gender induces speakers to highlight gender distinctions even in situations in which gender may not play an intrinsic role. Moreover, linguists have argued that the presence of “masculine generics” in some languages (e.g., the use of the noun “adam” to describe “man” as well as “human being” in the Hebrew and Turkish languages) has reinforced gender biases in the course of human history, equating “maleness and humanness” (13, p.169). Thus, perhaps not surprisingly, linguists as well as other scholars have persistently argued that gender biases have been reinforced by languages characterized by sex-based grammatical gender systems (11–13).

The study explores the geographical origins of sex-based grammatical gender based upon the existence of this language structure in 217 contemporary languages, originated from 74 lan-

⁴The use of this coding was designed to keep the signs of estimates across specifications aligned and simplify the interpretation.

guage families (63).⁵ Figure S1.3 and Table S10.2 describe the prevalence and the spatial distribution of sex-based grammatical gender systems in the dataset, demonstrating wide regional and cross-regional variations in the existence of these language structures. In particular, 37% of the languages in the sample are characterized by the presence of a sex-based grammatical gender system.

To further explore the proposed hypothesis, the analysis examines two alternative aspects of grammatical gender: (i) the existence of gender distinctions in independent personal pronouns (64), and (ii) the intensity of the grammatical gender system (combining information on the number of genders, the existence of sex-based grammatical gender, the system of gender assignment, and the existence of gender distinctions in independent personal pronouns) based on (63,65,66) and (64). As established in Appendix S8, the qualitative findings are independent of the measure used.

Figure S1.3: The Global Distribution of Sex-Based Grammatical Gender System

S1.A.3 Politeness Distinctions in Pronouns

Languages differ in the existence and the form of politeness distinctions. In particular, politeness distinctions in pronouns are present in some languages (e.g., the differential use of “tu” and

⁵While the WALS classifies the existence of grammatical gender for 245 contemporary languages, the classification of sex-based grammatical gender that is relevant for the main analysis restricts the sample to only 217 of them. In particular, the main analysis distinguishes between languages that have sex-based grammatical gender and those that do not have gender. Appendix S8 extends the analysis to the full sample by distinguishing between languages that have sex-based grammatical gender and those that do not have sex-based grammatical gender irrespective of whether they have grammatical gender.

“usted” in the Spanish language, “Du” and “Sie” in German, and “tu” and “vous” in French) while being absent from others. Moreover, linguists have underlined that the differential use of these pronouns is associated with the speaker’s social class and is closely associated with power distance (14).⁶

The study explores the geographical origins of politeness distinctions in pronouns based on the existence of second-person politeness distinctions in 207 contemporary languages, originated from 69 language families (71).⁷ Figure S1.4 and Table S10.3 describe the prevalence and the spatial distribution of politeness distinctions in the dataset, demonstrating wide regional and cross-regional variations in the existence of these language structures. In particular, 34% of the languages in the sample are characterized by the presence of politeness distinctions. Moreover, to further explore the proposed hypothesis, the analysis considers two alternative specifications of politeness distinctions: (i) the degree (rather than the existence) of politeness distinctions, and (ii) the existence of a high degree of politeness distinctions.⁸ As established in Appendix S9, the qualitative findings are independent of the classification used.

Figure S1.4: The Global Distribution of Politeness Distinctions

⁶In some languages, earlier linguistic forms predated politeness distinctions. In particular, in the Spanish language “usted” evolved from “vuestra merced” – your grace (68–70).

⁷ (71) classifies politeness distinctions, as reflected in second person pronouns, into four categories: (i) no politeness distinction, (ii) binary politeness distinction, (iii) multiple politeness distinctions, and (iv) dominantly avoided for politeness reasons. Hence, politeness distinctions are present if either one of the categories (ii)-(iv) are present.

⁸Using the classification of (71), the degree of politeness distinctions are captured by its assigned category, ranging from the lowest category – (i) to the highest by category – (iv). Similarly, high degree of politeness distinctions are defined as being associated with categories (iii) or (iv).

S1.B Main Independent Variables

In view of the hypothesized geographical origins of the explored language structures (i.e., the structure of the future tense, the presence of sex-based grammatical gender, and the existence of politeness distinctions), this subsection introduces the data on cross-regional variations in the hypothesized geographical determinants of these structures: (i) the historical levels of potential crop return, (ii) the historical levels of potential land suitability for the use of the plow, and (iii) the extent of ecological diversity.

Figure S1.5: The Global Distribution of Agricultural Determinants

S1.B.1 Crop Return in the Linguistic Homeland and the Urheimat

The potential historical crop return in the contemporary homeland of each language, as well as in its Urheimat (i.e., the homeland of the proto-language from which the language descended), is used in order to assess the geographical determinants of the structure of the future tense.

The potential historical crop return in the homeland and the Urheimat of a language (i.e., the potential return to crops that were native to these locations in the pre-1500 period, prior to the Columbian exchange) is measured, following the methodology advanced by (4), by the potential daily calories that could be generated by cultivating the crops that would maximize potential caloric yield in each cell of the homeland or the Urheimat, in the pre-1500CE era, given the crop growth cycle. The estimates of potential crop yield (measured in tons, per hectare, per year) and the corresponding crop growth cycle (measured as the number of days elapsed from planting to harvesting), for each $5' \times 5'$ cell in a global grid (i.e., approximately 100 square kilometers), are provided by the Global Agro-Ecological Zones (GAEZ) project of the Food and Agriculture Organization (FAO). These measures are based on agro-climatic estimates, under low level of inputs and rain-fed agriculture, capturing conditions that prevailed in early stages of development.⁹ The conversion of these crop yields into caloric yield (measured in millions of kilo calories, per hectare, per year), produces comparable estimates of crop yields, permitting the identification of the caloric maximizing crop in each cell, and given the crop growth cycle, the caloric daily return in the cell.¹⁰

These estimates have several virtues in identifying the effect of crop return in the contemporary homeland of a language or its Urheimat on the structure of the future tense. First, the focus on the potential crop return, rather than the actual one, assures that the effect of the structure of the future tense (and the associated time preference) on the choice of technology has no impact of the estimated crop return. Second, the use of the agro-climatic estimates, which are arguably orthogonal to human intervention, mitigates potential concerns about the exogeneity of these

⁹For each crop, GAEZ provides estimates for crop yield based on three alternative levels of inputs – high, medium, and low - and two possible sources of water supply – rain-fed and irrigation. Moreover, for each input-water source category, it provides two separate estimates for crop yield, based on agro-climatic conditions, that are arguably unaffected by human intervention, and agro-ecological constraints, that could potentially reflect human intervention. The FAO dataset provides for each cell in the agro-climatic grid the potential yield for each crop. These estimates account for the effect of temperature and moisture on the growth of the crop, the impact of pests, diseases and weeds on the yield, as well as climatic related “workability constraints”. In addition, each cell provides estimates of the growth cycle for each crop, capturing the days elapsed from the planting to harvesting.

¹⁰The caloric content of each crop is based on estimates of the United States Department of Agriculture Nutrient Database for Standard Reference.

estimates. Third, the use of estimates based on low level of inputs and rain-fed agriculture assures that the level of agricultural inputs, the irrigation method, and soil quality, do not reflect endogenous choices that could be potentially correlated with the structure of the future tense.

Hence, the average pre-1500CE potential daily caloric return (per hectare), R_ℓ , in the homeland or in the Urheimat of a language, ℓ , is

$$R_\ell = \frac{1}{|C_\ell|} \sum_{c \in C_\ell} \left(\frac{y_c}{g_c} \right), \quad (1)$$

where y_c is the cell specific pre-1500CE potential caloric yield of the crop that maximizes caloric output in cell c , g_c is this crop's growth cycle, C_ℓ is the set of cells in the language's homeland or Urheimat, and $|C_\ell|$ is the cardinality of this set. Figure S1.6 depicts the global distribution of the pre-1500CE potential crop return at the cell level.

Figure S1.6: The Global Distribution of Potential Crop Return (Pre-1500CE)

S1.B.2 Geographical Determinants of Plow Usage in the Linguistic Homeland and Urheimat

The potential historical suitability of the land for the usage of the plow in the homeland of each language, as well as in its Urheimat (i.e., the homeland of the proto-language from which the language descended), is used in order to assess the geographical determinants of sex-based grammatical gender.

The potential historical suitability of the land for the usage of the plow can be captured by either the average potential caloric yield attainable by crops that are suitable for the use of the plow in the pre-1500CE period (i.e., plow positive crops), or those that are unsuitable for the use

of the plow in the pre-1500CE period (i.e., plow negative crops), as classified by (20).¹¹ Since plow positive crops are characterized by limited regional variation (e.g., they are entirely absent from the Americas in the pre-1500 period), the analysis will focus on the adverse effect of the average suitability of plow negative crops on the adoption of the plow, as well as the positive effect of the average suitability of all crops on the desirability of the adoption of agricultural technology such as the plow.¹²

Figure S1.7: The Global Distribution of Potential pre-1500CE Caloric Yield (per Hectare) Across Plow Negative Crops

The average potential pre-1500CE caloric yield (per hectare) across plow negative crops in the homeland or in the Urheimat of a language, ℓ , P_ℓ , is

$$P_\ell = \frac{1}{|C_\ell|} \sum_{c \in C_\ell} y_c, \quad (2)$$

where y_c is the average potential pre-1500CE caloric yield of plow negative crops in cell c , C_ℓ is the set of cells in the language's homeland or Urheimat, and $|C_\ell|$ is the cardinality of this set. Figure S1.7 depicts the global distribution of the potential pre-1500CE average caloric yield (per hectare) across plow negative crops.

Similarly, the average potential pre-1500CE caloric yield (per hectare) across all crops in

¹¹Plow positive crops include wheat, barley, rye, buckwheat, teff, and wet rice, while plow negative crops include grains (such as millet, sorghum, dry rice, and maize), as well as all root and tree crops. Measures of the (non-caloric) suitability of the land for the use of the plow were first introduced by (3). The conversion of the crop yields of plow negative and plow positive crops into caloric yield (measured in millions of kilo calories, per hectare, per year), produces comparable estimates of crop yields and permits the construction of an index of soil suitability for the usage of the plow.

¹²In particular, since plow positive crops are absent from the Americas this measure cannot capture any of the existing variation in sex-based grammatical gender in these continents.

Figure S1.8: The Global Distribution of Potential pre-1500CE Caloric Yield (per Hectare) Across All Crops

the homeland or in the Urheimat of a language, ℓ , A_ℓ , is

$$A_\ell = \frac{1}{|C_\ell|} \sum_{c \in C_\ell} y_c, \quad (3)$$

where y_c is the average potential pre-1500CE caloric yield of all crops in cell c , C_ℓ is the set of cells in the language's homeland or Urheimat, and $|C_\ell|$ is the cardinality of this set. Figure S1.8 depicts the global distribution of the average potential pre-1500CE caloric yield (per hectare) across all crops.

S1.B.3 Ecological Diversity in the Linguistic Homeland and Urheimat

The extent of ecological diversity in the homeland of each language, as well as in its Urheimat (i.e., the homeland of the proto-language from which the language descended), is used in order to assess the geographical determinants of politeness distinctions.

The extent of ecological diversity, following (18), is measured as a Herfindahl index of the share of each territory that is occupied by different ecological zones. In particular, ecological diversity, E_ℓ , in the homeland or in the Urheimat of a language, ℓ , is

$$E_\ell = 1 - \sum_{j=1}^{16} (\theta_{\ell j})^2 \quad (4)$$

where $\theta_{\ell j}$ is the share of the homeland or the Urheimat of language ℓ that is located in ecological zone j , $j = 1, \dots, 16$.¹³

¹³ (76) provide a global dataset of biomes with 16 ecological zones: Boreal Forests/Taiga; Deserts and Xeric

S1.C Potentially Confounding Factors

The exploration of the geographical determinants of language structures accounts for the potentially confounding effects of a wide range of geographical factors, in the homeland of each language as well as in its *Urheimat* (i.e., the homeland of the proto-language from which the language descended). These factors may have affected the evolution of the structure of the future tense, sex-based grammatical gender, and politeness distinctions. In particular, it accounts for absolute latitude, average elevation, terrain ruggedness, coast length, as well as climatic conditions captured by the average, standard deviation, volatility, and spatial correlation of temperature and precipitation.¹⁴

The analysis further accounts for the length of the period in which no yield could be expected from the feasible set of crops (i.e., the expected number of days elapsed between the last potential harvest and the subsequent one). This unproductive period may have been central in the formation of cultural and linguistic characteristics since it had conceivably affected: (i) planning, intertemporal decisions, and therefore long-term orientation, (ii) the intensity of the use of the plow during the year as a whole and its implications for the emergence of gender roles, and (iii) cooperation in the development of storage technologies and its impact on the emergence of social hierarchy.

S1.D Cultural Measures

S1.D.1 Country-Level Measures

- **Language variables:** The three language variables (periphrastic future tense, sex-based grammatical gender and politeness distinctions) are measured as the share of the population who speaks a language with the the specific language structure. Hence, each of these variable is continuous between 0 and 1. Since the WALS does not include all languages in the world, the analysis is restricted only to the countries for which there is available data on at least 40% of the population. Data on the size of the population who speaks each language is taken from The Ethnologue.

Shrublands; Flooded Grasslands and Savannas; Inland Water; Mangroves; Mediterranean Forests, Woodlands and Scrub; Montane Grasslands and Shrublands; Rock and Ice; Temperate Broadleaf and Mixed Forests; Temperate Conifer Forests; Temperate Grasslands, Savannas and Shrublands; Tropical and Subtropical Coniferous Forests; Tropical and Subtropical Dry Broadleaf Forests; Tropical and Subtropical Grasslands, Savannas and Shrublands; Tropical and Subtropical Moist Broadleaf Forests; Tundra.

¹⁴Table S10.4 provides description of the variables used in the analysis and their summary statistics.

- Gender bias
 - **Female Labor Force Participation:** Labor force participation rate is the proportion of the population ages 15-64 that is economically active: all people who supply labor for the production of goods and services during a specified period. Year: 2003. The data were downloaded from the World Bank Gender Indicators.
 - **Share of Firms with Female Top Managers:** The percentage of firms in the private sector who have females as top managers. Top manager refers to the highest ranking manager or CEO of the establishment. Year: 2014. Taken from World Bank, Enterprise Surveys (<http://www.enterprisesurveys.org/>).
 - **Share of Full-Time Female Production Workers:** The percentage of permanent full time workers that are female, year: 2014. The data are downloaded from the World Bank, Enterprise Surveys, (<http://www.enterprisesurveys.org/>).
- Hierarchical bias
 - **Power Distance:** Taken from (77), which defines the measure as “the extent to which the less powerful members of organizations and institutions (like the family) accept and expect that power is distributed unequally.”
 - **Polity:** This measure is taken from The Polity IV Project (78), which “examines concomitant qualities of democratic and autocratic authority in governing institutions”. We use data for all years between 1800 and 1960.
 - **History of Statehood:** This measure represents the time elapsed since the beginning of historical statehood in modern countries. It is taken from (79).
- Long-term Orientation
 - **Long-Term Orientation:** This measure is taken from (77), which defines this measure as “the cultural value that stands for the fostering of virtues oriented toward future rewards, perseverance and thrift.”
 - **Patience:** This measure is taken from the Global Preference Survey (80, 81), a project run by the Briq institute. The measure is derived from the combination of responses to several survey measures, in which respondents choose between hypothetical binary choices between immediate and delayed financial rewards.

- **Education:** This measure is the average years of schooling in different countries in the world in 2010, as taken from (82).

S1.D.2 Individual level measures

Individual level measures are taken from the World Values Survey. For ease of interpretation, in some questions the answers were reverse coded. The measures include the following:

- Gender Bias
 - **Gender Bias:** is the first principal component to the answer of the following questions in the World Values Survey (WVS): (i) “When jobs are scarce men should have more right to a job than women”, (ii) “Men make better political leaders than women do”, (iii) “Women should have the same rights as men”, (iv) “A job is the best way for women to be independent”, (v) “It is a problem if women have more income than husband”, and (vi) “Men make better political leaders than women do” (questions c001, d059, e233, d063, d066, d066_b, d078).
 - **Ever married:** A dummy variables, based on answers to question x007.
 - **Ever Children:** A dummy variable assigned 1 if the answer to the question on the number of children (x011) is greater than zero.
 - **Number of Children:** The number of children as answered to question x011.
- Hierarchical Orientation
 - **Hierarchical Orientation:** This measure is an inverse coding of the Autonomy Index introduced in (83). It is based on the following questions in the World Values Survey (WVS): “Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important?” Individuals are considered to have higher hierarchical orientation if they answered “Obedience” or “Religious Faith” and lower if they answered “Independence” or “Determination, Perseverance” (y003, y020, y021).
 - **Respect for Parents:** The measure is based on answers to question a025, which states whether respect for parents should exist always, earned or neither.
 - **Favor Authority:** Based on question e018, which asks if a future change that will yield more respect for authority is desirable.

- **Freedom of Choice:** This measure captures how much freedom of choice and control over their own lives respondents feel they have. It is based on answers to question a173.
 - **No Participation in Political Protest:** This measure takes the value 1 if the respondent answered “No” to the following questions: If she signs petitions (e025), if she joins in boycotts (e026) and if she attends lawful demonstrations (e027).
 - **Favor Government Intervention:** This measure is based on the following questions: e035 (Income inequality should be more equalized or increased), e036 (Private or Government ownership of business should be increased), e037 (People vs. the government should take more responsibility) and e039 (Competition is good or harmful).
- Long-Term Orientation
 - **Long-Term Orientation:** coded as 1 if individual replied “Thrift, saving money and things” to question to question a038.
 - **Highest Completed Education Level:** The highest educational level obtained, as answered to question x025.
 - **Non-Vocational High School Attendance:** A dummy variable which indicates that the respondent has at least a non-vocational high school education, as reported in question x025.
 - **College Attendance:** A dummy variable which states if the respondent has at least a college degree, as reported in question x025.
 - **Income:** The income decile of the respondent, as reported in question x047.
 - **Savings:** A dummy variable whether the respondent’s family save during the last year prior to the interview, as reported in question x044.

S2 Complementarity between Culture and Linguistic Traits

This section, in line with the proposed hypothesis, establishes that cultural traits that are fundamental to the process of development are associated with complementary linguistic traits. In particular, there exists a highly significant positive association, across language groups, between (i) the presence of sex-based grammatical gender and the prevalence of gender bias in society (e.g., beliefs about gender roles, gender gaps in labor force participation, occupation, and education; Tables S2.1, S2.2, S2.3, S2.4), (ii) the existence of politeness distinctions and the pervasiveness of hierarchical orientation (e.g., beliefs about the role of authority and government, respect for authority and parents, power distance, autocracy; Tables S2.1, S2.2, S2.5), and (iii) the existence of periphrastic future tense and degree of long-term orientation in society (e.g., patience, human capital formation, savings, income; Tables S2.1, S2.2, S2.6).

Table S2.1: Association Between Cultural and Linguistic Traits
Language-level Analysis

	Cultural Values					
	Long-Term Orientation		Gender Bias		Hierarchy Index	
	All	Old World	All	Old World	All	Old World
	(1)	(2)	(3)	(4)	(5)	(6)
Periphrastic Future Tense	0.15*** (0.04)	0.17*** (0.05)				
Existence of Sex-Based Grammatical Gender			0.42** (0.16)	0.18* (0.09)		
Politeness Distinctions					0.23** (0.09)	0.29*** (0.06)
Caloric Crop Return	0.29* (0.16)	0.26 (0.18)				
Caloric Plow Suitability			-0.16 (0.43)	0.25 (0.29)		
Ecological Diversity					0.14 (0.13)	0.05 (0.14)
Regional FE	Yes	Yes	No	No	No	No
Main Geographical Controls	Yes	Yes	Yes	Yes	Yes	Yes
Average Caloric Yield	No	No	Yes	Yes	Yes	Yes
Adjusted- R^2	0.24	0.26	0.33	0.38	0.27	0.36
Observations	106	100	33	30	53	50

Notes: This table documents the statistically and economically significant positive association between cultural traits and linguistic structures. The results account for the geographical origin of the periphrastic future tense and other geographical characteristics of the contemporary homeland of the language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S2.2: Contemporary Culture and Linguistic Traits
Country-Level Analysis

Socio-Economic Outcomes								
Female Labor Force Participation	Gender Bias		Hierarchy			Long-Term Orientation		
	Share of Firms with Top Female Managers	Female Share of Full Time Production Workers	Power Distance	State History	Degree of Democracy	LTO	Patience	Education
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sex-Based Gender	-0.29** (0.14)	-0.36* (0.18)	-0.53*** (0.15)					
Politeness Distinctions			0.29** (0.13)	0.38*** (0.13)	-0.44*** (0.13)			
Periphrastic Future						0.35*** (0.11)	0.42*** (0.13)	0.47*** (0.09)
Adjusted- R^2	0.07	0.11	0.27	0.13	0.18	0.11	0.17	0.21
Observations	60	41	39	78	60	73	58	73

Notes: This table shows the statistically and economically significant association between contemporary culture and linguistic traits. Specifically, it shows the negative correlation between the share of people in a country who speak a language with sex-based grammatical gender and they country's (i) female labor force participation; (ii) share of firms with a top female manager; and (iii) female share of full time production workers. It shows the positive correlation between the share of people who speak languages characterized by politeness distinctions and the country's level of (i) power distance; and (ii) the duration of statehood, and a negative correlation between the share of people who speak languages characterized by politeness distinctions and the country's level of democracy. Finally, it shows the positive correlation of the share of people who speak languages characterized by a periphrastic future tense and the country's (i) long-term orientation; (ii) Patience; and (iii) Average years of schooling. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests; All regressions include a constant.

Table S2.3: Association Between Sex-Based Grammatical Gender and Gender Bias
Individual-level Analysis

	Gender Bias				
	Ever Married	Ever Children	Number of Children	Housewife	Main Income Earner
	(1)	(2)	(3)	(4)	(5)
Sex-Based Grammatical Gender	0.075** (0.034) [0.034]	0.059* (0.030) [0.036]	0.327*** (0.084) [0.161]	0.108 (0.089) [0.086]	-0.039** (0.016) [0.022]
Age	0.020*** (0.002) [0.001]	0.017*** (0.002) [0.001]	0.082*** (0.009) [0.008]	0.007*** (0.002) [0.002]	0.005*** (0.001) [0.001]
Sex-Based Grammatical Gender × Age	-0.001** (0.000) [0.001]	-0.001*** (0.000) [0.001]	-0.011* (0.006) [0.007]	-0.003** (0.001) [0.001]	0.002* (0.001) [0.001]
Country FE	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes
Gender FE	Yes	Yes	Yes	Yes	Yes
Mean Dependent Variable	0.79	0.79	2.06	0.30	0.27
Adjusted- R^2	0.35	0.36	0.38	0.30	0.15
R^2	0.36	0.37	0.39	0.30	0.17
Observations	117731	115899	97431	116622	90270

Notes: This table establishes the statistically and economically significant association of speaking a language characterized by sex-based grammatical gender and several individual-level economic outcomes. In particular, it shows that women who speak a language with sex-based grammatical gender have a higher probability to marry, have more children and become a housewife, and to do so at a younger age (than those who speak a language without sex-based grammatical gender). Furthermore, they have a lower probability to become the main income earners, and they will not be the main income earners for a longer period. All regressions include country, age, year and gender fixed effect. Heteroskedasticity robust standard error estimates clustered at the country and language family levels are reported in parentheses and at the country and language levels in brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S2.4: Association Between Intensity of Grammatical Gender and Gender Bias
Individual-level Analysis

	Gender Bias				
	Ever Married	Ever Children	Number of Children	Housewife	Main Income Earner
	(1)	(2)	(3)	(4)	(5)
Intensity of Grammatical Gender	0.065** (0.025) [0.035]	0.042* (0.020) [0.032]	0.096 (0.077) [0.121]	0.080** (0.034) [0.041]	-0.017** (0.008) [0.007]
Age	0.020*** (0.002) [0.001]	0.017*** (0.002) [0.002]	0.073*** (0.009) [0.007]	0.006*** (0.002) [0.001]	0.006*** (0.001) [0.001]
Intensity of Grammatical Gender × Age	-0.001** (0.000) [0.001]	-0.001* (0.000) [0.001]	0.000 (0.002) [0.002]	-0.001*** (0.000) [0.001]	0.000 (0.000) [0.000]
Country FE	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes
Gender FE	Yes	Yes	Yes	Yes	Yes
Mean Dependent Variable	0.78	0.79	2.05	0.29	0.27
Adjusted- R^2	0.35	0.36	0.38	0.29	0.15
Observations	111553	110277	93550	110447	85981

Notes: This table establishes the statistically and economically significant association between speaking a language with a higher intensity of grammatical gender and several individual-level economic outcomes. In particular, it shows that women who speak a language with a higher intensity of grammatical gender have a higher probability to marry, have more children and become housewives, and to do so at a younger age (than those who speak a language with a low intensity of grammatical gender). Furthermore, they have a lower probability to become the main income earners, and they will not be the main income earners for a longer period. All regressions include country, age, year and gender fixed effect. Heteroskedasticity robust standard error estimates clustered at the country and language family levels are reported in parentheses and at the country and language levels in brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S2.5: Association Between Politeness Distinctions and Hierarchical Orientation
Individual-level Analysis

	Hierarchical Orientation				
	Respect for Parents	Favor Authority	Freedom of Choice	No Participation in Political Protest	Favor Government Intervention
	(1)	(2)	(3)	(4)	(5)
Politeness Distinctions	0.034** (0.012) [0.015]	0.021* (0.010) [0.012]	-0.037*** (0.011) [0.016]	0.073*** (0.005) [0.013]	0.084*** (0.022) [0.051]
Country FE	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes
Gender FE	Yes	Yes	Yes	Yes	Yes
Mean Dependent Variable	1.81	0.61	0.71	0.69	-0.26
Adjusted- R^2	0.13	0.18	0.09	0.15	0.37
Observations	132523	253614	259567	248801	259353

Notes: This table establishes the statistically and economically association between speaking a language with politeness distinctions and several outcomes that reflect hierarchical orientation. In particular, the table shows that speaking a language with politeness distinctions is positively correlated with having respect for parents, favouring authority and governments intervention. Furthermore, it shows that speaking a language with politeness distinctions is negatively associated with favouring freedom of choice and participation in political protests. All regressions include country, year, age and gender fixed effects. Heteroskedasticity robust standard error estimates clustered at the country and language family levels are reported in parentheses and at the country and language levels in brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S2.6: Association Between Periphrastic Future Tense and Long-Term Orientation
Individual-level Analysis

	Long-Term Orientation				
	Highest Com- plemented Education Level	Non- Vocational High School Attendance	College Attendance	Income	Savings
	(1)	(2)	(3)	(4)	(5)
Periphrastic Future Tense	0.452*** (0.100) [0.165]	0.078*** (0.019) [0.033]	0.063*** (0.008) [0.023]	0.595*** (0.170) [0.244]	0.035*** (0.009) [0.016]
Country FE	Yes	Yes	Yes	Yes	Yes
Wave FE	Yes	Yes	Yes	Yes	Yes
Age FE	Yes	Yes	Yes	Yes	Yes
Gender FE	Yes	Yes	Yes	Yes	Yes
Mean Dependent Variable	4.77	0.43	0.25	4.64	0.20
Adjusted- R^2	0.23	0.17	0.13	0.18	0.16
Observations	315898	318748	318748	252545	310274

Notes: This table establishes the statistically and economically positive association between speaking a language with a periphrastic future tense and long-term oriented economic outcomes. In particular, this table shows the speaking a language with a periphrastic future tense is associated with higher levels of education attainment, a higher income and a higher probability to save. All regressions have country, year, age and gender fixed effects. Heteroskedasticity robust standard error estimates clustered at the country and language family levels are reported in parentheses and at the country and language levels in brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Figure S2.1: Contemporary Cultural & Linguistic Traits

The figure depicts the association between linguistic and cultural traits across languages as reflected in socio-economic outcomes and preferences.

S3 The Structure of the Empirical Analysis

This section presents the empirical strategy for the exploration of the relation between language structures and their hypothesized geographical determinants: (i) the historical levels of potential crop return to agricultural investment, in the language’s homeland and its Urheimat, and the existence of periphrastic future tense, (ii) the historical levels of potential land suitability for plow-based agriculture, in the language’s homeland and its Urheimat, and the existence of sex-based grammatical gender, and (iii) the extent of ecological diversity, in the language’s homeland and its Urheimat, and the existence of politeness distinctions.

S3.A The Empirical Model

The empirical analysis estimates the association between the hypothesized geographical determinants and linguistic traits using a linear probability model, given the binary nature of the dependent variables. In particular, it estimates the following empirical model using ordinary least squares (OLS):

$$S_\ell = \beta_0 + \beta_1 D_{\ell S} + \sum_j \gamma_{0j} X_{\ell j} + \sum_c \gamma_c \delta_{\ell c} + \varepsilon_\ell, \quad (5)$$

where S_ℓ indicates whether structure S =(periphrastic future tense, sex-based grammatical gender, politeness distinctions) exists in language ℓ , $D_{\ell S}$ is the hypothesized geographical determinant of structure S in the homeland of language ℓ and in its Urheimat (i.e., pre-1500CE potential crop return, R_ℓ , the average potential pre-1500CE caloric yield of plow negative crops, P_ℓ , and all crops, A_ℓ , and the extent of ecological diversity, E_ℓ), $\{X_{\ell j}\}$ is a set of potentially confounding geographical characteristics in the homeland of language ℓ and in its Urheimat, $\{\delta_{\ell c}\}$ is a complete set of regional fixed-effects in the homeland of language ℓ and in its Urheimat, and ε_ℓ is a language-specific error term.

Moreover, given the binary coding of each language structure, an alternative specification is considered and estimated using a Probit model. As reported in the Appendix, this alternative estimation model generates qualitatively similar results.

S3.B The Empirical Strategy

The empirical analysis of the geographical determinants of language structures advances several strategies to mitigate potential concerns regarding the role of reverse causality, omitted variables, and sorting in the observed association between: (i) variations in geographical characteristics that were conducive to higher natural return to agricultural investment and the existing cross-language variations in the structure of the future tense, (ii) agricultural determinants of gender gap in agricultural productivity and the existence of sex-based grammatical gender, and (iii) the ecological origins of hierarchical societies and the emergence of politeness distinctions.

First, the observed associations between the hypothesized geographical determinants of linguistic traits and observed language structures may reflect reverse causality from language structures, and their associated cultural traits, to these geographical attributes. In particular, the conceivable effect of periphrastic future tense, and the associated long-term orientation, on the choice of agricultural technologies, may affect crop returns. Moreover, sex-based grammatical gender, and the associated gender bias, may have conceivably affected the adoption of the plow. Hence, in order to overcome these potential concerns about reverse causality, the empirical analysis exploits variations in potential crop return (associated with agro-climatic conditions that are orthogonal to human intervention), rather than the, potentially endogenous, actual crop return. Furthermore, the analysis is based upon the potential use of the plow (as captured by potential agro-climatic yield of crops suitable for the usage of the plow), rather than the actual usage.¹⁵

Second, the associations between geographical and linguistic traits may be governed by omitted institutional, cultural, geographical and human characteristics. Thus, in order to mitigate these concerns, the empirical analysis mirrors the epidemiological approach to cultural diffusion and identifies the geographical origins of linguistic traits based on variation across languages located outside their *Urheimat* (i.e., outside the ancestral homeland of their proto-language), accounting for the host region fixed-effects and thus capturing unobserved time-invariant heterogeneity at the regional level. Moreover, the analysis accounts for the potentially confounding effects of a wide range of geographical factors, in the homeland of each language, as well as in its *Urheimat*, factors which may have affected the evolution of the structure of the future tense, sex-based grammatical gender, and politeness distinctions (e.g., absolute latitude,

¹⁵Instrumenting for the actual usage of the plow with the suitability of the land for the usage of the plow (as established in Figure 4 and Supplemental MaterialS5.C), would reduce the sample considerably, given the limited data on the actual usage of the plow, and additionally, the analysis would violate the exclusion restriction.

average elevation, terrain ruggedness, coast length, as well as climatic conditions captured by the average, standard deviation, volatility, and spatial correlation of temperature and precipitation). Furthermore, the empirical analysis accounts for spatial auto-correlation and explores whether selection on unobservables is significant.¹⁶

Third, the observed associations between geographical and linguistic traits may reflect the sorting of individuals into geographical niches that complement their linguistic and cultural traits. While sorting would not affect the nature of the association between these geographical and linguistic characteristics (i.e., variations in geographical characteristics across the globe would still be the origin of cross-language variation in linguistic traits), it would weaken the cultural interpretation of the underlying mechanism. Hence, to alleviate concerns about the potential sorting of individuals who use periphrastic future tense into regions with high crop return, the analysis explores whether potential crop return in the Urheimat of a language, rather than its contemporary homeland, is associated with the presence periphrastic future tense. Similarly, to mitigate concerns about the potential sorting of individuals who use sex-based grammatical gender into regions suitable for the usage of the plow, the analysis explores whether the suitability of land for the usage of the plow in the Urheimat, rather than in the contemporary homeland of daughter languages, is associated with sex-based grammatical gender.¹⁷

Fourth, the hypothesized associations between geographical and linguistic traits may reflect a broader association between each of these geographical factors and a wide range of linguistic traits, rather than a unique association between each of these geographical characteristics and its hypothesized corresponding linguistic trait, weakening the proposed hypothesis. Hence, in order to overcome this potential concern, the empirical analysis explores whether alternative language structures (as reported by WALS), such as the existence of other temporal structures (e.g., the past tense and the perfect tense), as well as non-temporal structures (e.g., possessive, evidentiality, and the number of consonants and colors), are associated with each of the proposed geographical roots.

Fifth, the proposed hypothesis suggests that the evolution of linguistic traits is a by-product

¹⁶It should be noted that since the available data for a large fraction of language families consists of only a few languages, a meaningful analysis based on variations in language structures within a linguistic family is precluded. Nevertheless, given the common ancestry of daughter languages, the analysis is based on standard errors estimates clustered at the language, language genus or language family levels. Specifically, the analysis based on the geographical characteristics of the Urheimat employs standard errors estimates clustered at the language family level.

¹⁷In light of the evidence about the greater adaptability of politeness distinctions to environmental and political changes (30–32, 41, 42), this strategy does not fully mitigate concerns on the role of sorting in the observed association between ecological diversity and politeness distinctions.

of the coevolution of cultural and linguistic traits in the course of human history, as governed by deeply rooted, culturally-embodied, common geographical roots. Hence, the proposed hypothesis rests upon the importance of the culturally-embodied effect of geographical characteristics. In order to explore the culturally-embodied effect, the empirical analysis examines whether the geographical characteristics in the Urheimat, rather than the homeland of daughter languages, are the significant factors that govern the evolution of language structures. Thus, it explores whether the culturally-embodied geographical determinants in the Urheimat, rather than the direct effect of geography (as would have been reflected by the significance of the geographical characteristics in the contemporary homeland of daughter languages), governed the evolution of linguistic traits.

S4 Geographical Origins of Periphrastic Future Tense

The hypothesized coevolution of long-term orientation and the structure of the future tense suggests that in societies characterized by higher long-term orientation in the distant past, periphrastic future tense, and its associated future orientation, have emerged and persisted over time, reinforcing existing long-term oriented behavior. Thus, in view of the documented positive effect of the natural return to agricultural investment, as captured by caloric crop return, on long-term orientation (4),¹⁸ crop return is hypothesized to be associated with the prevalence of periphrastic future tense.¹⁹

This section analyzes the relation between the existence of periphrastic future tense across contemporary languages and the historical levels of potential crop return in their contemporary geographical homelands (i.e., the indigenous regions that spoke the language in the pre-colonial as well as in the post-colonial era), as well as in their ancestral homeland (i.e., their Urheimat).

¹⁸See also, (89) and (90).

¹⁹The ability to delay gratification is relevant for agricultural as well as non-agricultural societies. However, in light of the importance attributed to the ability to delay gratification in the cultivation of crops and in view of data availability, the analysis is centered around those geographical characteristics that were conducive for higher caloric return to agricultural investment and establish their impact on the prevalence of long-term orientation in society and the existence of periphrastic future tense. There does not exist extensive data on exogenous geographical sources of variation in caloric yield from other food sources that would enable the analysis to document this complementary channel. However, it should be noted that evidence suggests that the transition to agriculture has a larger impact in the level of patience across societies and thus potentially on the emergence of periphrastic future tense. For instance, experiments among the Tsimane suggest that among this tribe those who transitioned to agriculture have higher levels of long-term orientation (46). Similarly, (47) show evidence that agriculturalists have higher long-term orientation than fishermen.

In particular, Table S4.1 explores the association between pre-1500CE crop return and the existence of periphrastic future tense.²⁰ As documented in column (1), the unconditional correlation between pre-1500CE crop return in the language's contemporary homeland and the existence of periphrastic future tense is positive and statistically significant at the 5% level, suggesting that a one standard deviation increase in crop return in the language's contemporary homeland is associated with a 6 percentage points increase in the probability that the language is characterized by a periphrastic future tense. Furthermore, as documented in column (2) the estimated relationship remains stable once the potentially confounding effects of the geographical characteristics (i.e., absolute latitude, mean elevation, terrain ruggedness, and the length of its sea coast) in the contemporary homeland of each language, are accounted for.

Moreover, since crop return is estimated based on agro-climatic conditions, its association with the structure of the future tense might be capturing the potentially direct impact of climatic conditions on the existence of periphrastic future. In particular, variability of climatic conditions may affect the riskiness of agricultural investment, and therefore the effective crop return, while their spatial correlation may affect the potential for spatial diversification of risk. Nevertheless, as established in columns (3) and (4), accounting for the potentially confounding effects of average temperature and precipitation, as well as their standard deviations, volatility, and spatial correlation, does not alter the qualitative results.

Furthermore, the length of the period within each year in which the land is not suitable for agriculture – the unproductive period – may have conceivably triggered planning (e.g., storage and intertemporal trade) and may have therefore reinforced long-term orientation and the existence of periphrastic future tense. Nevertheless, as indicated by the estimates in column (5), accounting for the length of the unproductive period, reinforces the findings and the coefficient on crop return increases and becomes highly statistically significant.²¹ Finally, as reported in column (6), accounting for regional fixed-effects, and therefore for unobserved time-invariant heterogeneity at the regional level, reinforces the magnitude and the statistical significance of the association between pre-1500CE crop return and the existence of periphrastic future tense. In particular, a one standard deviation increase in pre-1500CE crop return in the language's contemporary homeland is associated with an 11 percentage points increase in the probability

²⁰Table S7.1 reports the estimated coefficients on the control variables in Table S4.1 .

²¹Interestingly, as shown in the Appendix, the length of the unproductive period, as well, is significantly associated with the probability of existence of the periphrastic future tense. However, unlike the crop return, the relation between the unproductive period and the existence of periphrastic future tense is not robust to the other specifications and its semi-partial R^2 is smaller than the one of crop return.

Table S4.1: Geographical Origins of the Future Tense (OLS)

	Existence of Future Tense					
	(1)	(2)	(3)	(4)	(5)	(6)
Crop Return (pre-1500CE)	0.06** (0.03)	0.06** (0.03)	0.07** (0.03)	0.07** (0.03)	0.09*** (0.03)	0.11*** (0.03)
Geographical Controls	No	Yes	Yes	Yes	Yes	Yes
Precipitation Controls	No	No	Yes	Yes	Yes	Yes
Temperature Controls	No	No	No	Yes	Yes	Yes
Unproductive Period	No	No	No	No	Yes	Yes
Regional FE	No	No	No	No	No	Yes
Adjusted- R^2	0.01	0.02	0.02	0.06	0.07	0.11
Observations	275	275	275	275	275	275

Notes: This table documents the positive, statistically, and economically significant association between a region pre-1500CE potential crop return and the existence of periphrastic future tense in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

that the language is characterized by a periphrastic future tense.

S4.A Robustness

The findings in Table S4.1 lend credence to the hypothesis that the natural return to agricultural investment in the language's contemporary homeland, and its effect on time preference and long-term oriented behavior, contributed to the emergence of periphrastic future tense. Nevertheless, the results conceivably might be biased due to omitted institutional, cultural, and human characteristics, precluding a causal interpretation of the estimated coefficients. While concerns about omitted variables will be directly addressed and mitigated in the next subsection, the remaining analysis in this section documents the robustness of the results to: (i) an alternative estimation method, (ii) spatial autocorrelation, and (iii) clustering of standard errors, while demonstrating that the findings are unlikely to be driven by selection on unobservables.

S4.A.1 Estimation Method, Spatial-Autocorrelation and Selection on Unobservables

Table S7.2 (Panel A) suggests that the estimated average marginal effects of the Probit model are nearly identical to the estimated associations using OLS. Furthermore, Table S7.2 (Panel B)

documents the robustness of the results to the clustering of the standard errors at the level of the language genus (i.e., a group of phylogenetically related languages inside a linguistic family), as well as spatial auto-correlation. In particular, the statistical significance is unaffected by the various methods used to mitigate the potential violation of the independence assumption.

Table S7.2 (Panel B) further explores the potential bias that could be generated by omitted variables. In particular, using statistics on selection on observables and unobservables (91, 92), it suggests that the degree of omitted variable bias is low and is unlikely to explain the magnitude of the estimated association between crop return and the existence of periphrastic future tense. In particular, omitted factors would need to be 2-3 times stronger and negatively correlated with crop return, in order to account for the estimated association between crop return and the emergence of periphrastic future tense. Thus, the estimated coefficients could be considered as the lower bound of the true association. The bias-adjusted estimated coefficient on pre-1500CE crop return is strictly positive and larger than the OLS estimate. In particular, as reported in column (6), following (92), the bias-corrected coefficient of crop return, assuming that the unobservables are as strongly correlated with pre-1500CE crop return as the set of observables that are accounted for, implies that a one standard deviation increase in crop return in the language's contemporary homeland is associated with a 13 percentage points increase in the probability of the existence of periphrastic future tense in the language.²²

S4.A.2 Crop Return and Other Language Structures

The associations between the return to agricultural investment in the language's contemporary homeland, as captured by pre-1500CE crop return, and the existence of periphrastic future tense may reflect a broader association between crop return and a wide range of linguistic traits, rather than a unique association with periphrastic future tense, weakening the proposed hypothesis. Moreover, the association between crop return in the language's homeland and the existence of periphrastic future tense in the language may reflect a broader association between a whole range of linguistic traits and long-term orientation.

Table S14.1 suggests, however, that alternative language structures (as reported by WALS), such as the existence of other temporal and non-temporal structures, are largely uncorrelated with the natural return to agricultural investment, suggesting that it is indeed uniquely associated

²²If the unobservables were assumed to explain all the remaining variation in the probability of existence of periphrastic future tense, then the estimated bias-adjusted impact of crop return would be even larger, implying that a one standard deviation increase in crop return in the language's contemporary homeland would be associated with 40 percentage points increase in the probability of existence of periphrastic future tense in the language.

with the structure of the future tense.²³ In particular, pre-1500 crop return is not significantly associated with the existence of either the past or the perfect tense, and it is orthogonal to non-temporal linguistic characteristics such as: (i) the existence of possessive classifications, (ii) the existence of coding for evidentiality, (iii) the number of consonants, and (iv) the number of colors.²⁴

To alleviate concerns about potential association with other inflectional structures or other tenses, we show in the new Table S14.2 for the common sample of languages that have data for all these language structures, only the periphrastic future tense is strongly associated with the geographical determinant of long-term orientation. In particular, notice that not only is the association with the periphrastic future tense statistically significant at the 1% (and insignificant for the others), but the estimated coefficient is 2.1-2.5 times the size of the estimate for the other structures. Moreover, the variation uniquely associated with crop return (i.e., its semi-partial R-squared) explains 6% of the total variation in the future tense, while the variation uniquely associated to each of all the other controls taken together (i.e., the sum of the semi-partial R-squared of all the other controls) explains 14% of the total variation in the future tense. Thus, crop return uniquely explains about 43% of the variation in future tense that can be explained by the other controls. On the other hand, the variation uniquely associated with crop return explains only 1-2% of the total variation in the other structures, which represents about 5-14% of the variation that can be explained by the other controls. This further suggests that the relation between crop return and periphrastic future tense is indeed unique. In addition, as shown in Table S7.4, there is no statistical association between the measure of long-term orientation and these other tense structures, supporting our hypothesis that the periphrastic future tense is complementary to long-term oriented behavior.

Furthermore, using all available data on language structures as reported in WALS, Figure S14.3 documents the orthogonality of crop return for nearly all language structures except for the structure of the future tense.²⁵ These results lend further support to the importance of the

²³This result reinforces earlier findings by (4) that pre-1500CE crop return only affects time preference and does not have a significant association with the other cultural traits.

²⁴While the association between crop return and the existence of possessive classification is marginally significant in the OLS regressions, it is insignificant if Probit is used.

²⁵Specifically, the only other structures that are statistically associated with crop return are: existence of politeness distinctions (variables 45A in WALS), existence of cases (49A), number of cases (50A), existence of inflectional tenses (69A), and existence of purpose clauses (125A). The association with politeness distinctions reflects the positive impact of agricultural productivity, and the associated surplus, with the emergence of hierarchical societies. Furthermore, the association with inflectional tenses reflects by construction the negative association between the existence of a periphrastic future tense and the existence of an inflectional tense. Finally, the association

association between pre-1500 crop return and the emergence of the periphrastic future tense.

S4.B Geographical Origins: Contemporary Homeland vs. Urheimat

This subsection exploits the descent of contemporary (daughter) languages from proto-languages to explore the relative contribution of the return to agricultural investment in (i) the contemporary homeland of the language and (ii) the ancestral homeland of the language (i.e., its Urheimat) to the existence of a periphrastic future tense in daughter languages. Specifically, the proposed hypothesis suggests that the evolution of linguistic traits is a by-product of the co-evolution of cultural and linguistic traits in the course of human history, as governed by deeply rooted, culturally-embodied, common geographical roots. In order to examine this hypothesis, the empirical analysis explores whether potential crop return in the Urheimat, rather than the homeland of daughter languages, is associated with the presence of periphrastic future tense. This setting permits the exploration of the role of culturally-embodied geographical determinants in the Urheimat, rather than the direct effect of geography in the evolution of linguistic traits.

Furthermore, this analysis mitigates concerns about the role of omitted variables in the associations between the return to agricultural investment and the existence of periphrastic future tense, while overcoming concerns about the potential sorting of individuals into geographical niches that complement their linguistic and cultural traits. Moreover, the analysis sheds light on the historical depth of the observed relationship between the return to agricultural investment and the existence of periphrastic future tense, as well as on the role of the culturally-embodied geographical characteristics in the Urheimat, rather than the direct effect of geography, in the evolution of periphrastic future tense.

In particular, in order to further mitigate concerns about the importance of omitted institutional, cultural, geographical and human characteristics in the associations between the return to agricultural investment and the existence of periphrastic future tense, the analysis mirrors the epidemiological approach to cultural diffusion, and explores the geographical origins of the existence of periphrastic future tense based on variation across languages located outside the Urheimat (i.e., the ancestral homeland of their proto-language). This approach permits the analysis to account for the host region fixed-effects and thus to capture unobserved time-invariant

with the remaining language structures reflects the common elements among certain aspects of language structures. For example, the existence of temporal case structures may be associated with the existence of tenses, and the future in particular (93).

heterogeneity at the regional level.

Moreover, alleviating concerns about the potential sorting of individuals who use periphrastic future tense into regions with high crop return, the analysis explores whether potential crop return in the Urheimat, rather than in the contemporary homeland of daughter languages, is associated with the presence periphrastic future tense. While sorting would not affect the nature of the association between these geographical and linguistic characteristics (i.e., variations in geographical characteristics across the globe would still be the origin of cross-language variation in linguistic traits), it would weaken the cultural interpretation of the underlying mechanism. Finally, the distinction between the impact of the geographical characteristics in the Urheimat and the contemporary homeland, permit the analysis to explore the historical depth of the observed relationship between the return to agricultural investment and the existence of periphrastic future tense.

Table S4.2: The Persistent Impact of Geographical Characteristics in the Urheimat, rather than in the Contemporary Linguistic Homeland, on the Existence of Periphrastic Future Tense: Daughter Languages Outside the Urheimat

	The Existence of Periphrastic Future Tense					
	Migratory Distance to the Urheimat					
	Any Distance			At Least 1 Week		
	Homeland	Urheimat	Both	Homeland	Urheimat	Both
	(1)	(2)	(3)	(4)	(5)	(6)
Homeland's Crop Return (pre-1500CE)	0.11*** (0.04)		0.01 (0.05)	0.12** (0.06)		-0.01 (0.06)
Urheimat's Crop Return (pre-1500CE)		0.16** (0.08)	0.17** (0.08)		0.23*** (0.08)	0.24** (0.09)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes
Homeland Geographical Characteristics	Yes	No	Yes	Yes	No	Yes
Urheimat Geographical Characteristics	No	Yes	Yes	No	Yes	Yes
Adjusted- R^2	0.10	0.13	0.17	0.07	0.12	0.15
Observations	255	255	255	195	195	195
Language Families	59	59	59	32	32	32

Notes: This table documents the importance of pre-1500CE crop return in the Urheimat, rather than the contemporary homeland, to the presence of periphrastic future tense in daughter languages located outside of the Urheimat, accounting for regional fixed-effects and other geographical characteristics. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length. Columns (1)-(3) consider daughter languages who are located anywhere outside of the Urheimat, while column (4)-(6) those who are at least a week of migratory distance from the Urheimat. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S4.2 explores the relative contributions of pre-1500CE crop return in the contemporary homeland of a language, in comparison to its ancestral homeland (i.e., its Urheimat), to the presence of periphrastic future tense in daughter languages. It suggests that the existence of periphrastic future tense among daughter languages located outside their Urheimat is governed predominantly by crop return in the Urheimat.²⁶ Considering daughter languages anywhere outside of the Urheimat, Columns (1) - (3) suggest that while crop return in the contemporary homeland is significantly associated with the existence of periphrastic future tense among migrant languages (column (1)), the association between the Urheimat's crop return and the existence of periphrastic future tense of daughter languages located outside of the Urheimat is much larger (column (2)). Moreover, considering jointly the contribution of crop return in the contemporary homeland and the Urheimat, the impact of the crop return in the Urheimat remains significant, while the impact of crop return in the contemporary homeland dissipates entirely, suggesting that that crop return in the Urheimat predominately influenced the existence of the future tense.²⁷ Furthermore, considering daughter languages at a migratory distance of one week from the Urheimat (i.e, approximately 1300km away from the Urheimat), and thus diminishing the association between the geographical characteristics in the contemporary homeland and the Urheimat, columns (3)-(6) reinforce the findings about the domination of the crop return in the Urheimat rather than in the contemporary homeland. In particular, accounting for regional fixed-effects and other geographical characteristics of the Urheimat, a one standard deviation increase in the crop return in the Urheimat is associated with a 21-25 percentage points increase in the probability that a migrant daughter language would have a periphrastic future tense.

The analysis suggests that periphrastic future tense began its formation in the distant past when the speakers of the proto-language still inhabited the Urheimat of the language, reflecting the portable culturally-embodied effect of crop return rather than the direct effect of these geographical characteristics.²⁸ The findings further suggest that the association is not driven

²⁶Given the potential dependence of languages within the same language family, standard errors are clustered at this level.

²⁷Given the greater uncertainty about the geographical extent of the Urheimat, as opposed to the contemporary linguistic homeland, measurement errors in the estimates of crop return in the Urheimat are likely to be larger than those from crop return in the contemporary homeland and a-priori the horse race between the two variables should have resulted in fact in the decline in the significance of crop return in the Urheimat.

²⁸While the major dichotomy between periphrastic and inflectional future tense has emerged in the Urheimat, more nuanced distinctions among languages characterized by periphrastic future tense may have evolved subsequently. For instance, the current form of periphrastic future tense in the English language emerged in the last millennium from another type of periphrastic future tense (closer in its structure to the future tense in Germanic

by either omitted factors or geographical sorting, and they reflect the deep-historical culturally-embodied origins of the structure of the future tense.

S4.C Mechanism

This section presents additional supportive evidence for the hypothesized geographical origins of the structure of the future tense. In view of the evidence about the association between crop return and long-term orientation (4), a positive effect of crop return on the emergence of periphrastic future tense would necessitate the use of agriculture as the main source of subsistence. In particular, as illustrated in Figure S4.1, the proposed hypothesis would imply that: (i) higher crop return ought to be associated with a greater intensity of agriculture, and (ii) greater intensity of agriculture ought to be associated with the existence of periphrastic future tense.

Figure S4.1: Crop Return, Agricultural Intensity & Future Tense

Table S4.3 presents supporting evidence for the proposed mechanism. In particular, it documents the robust positive association between crop return and the level of agricultural intensity in a pre-colonial society, which in turn is positively associated with the emergence of the periphrastic future tense. Using data from the Ethnographic Atlas (51) on the importance of patterns of subsistence – hunting, gathering, fishing, animal husbandry and crop cultivation – the analysis explores the association between crop return and agricultural intensity (i.e., the level of dependence on agriculture) across ethnic groups, reflecting the effective crop return in the ethnic homeland. As suggested by Columns (1) and (2) in Table S4.3 societies inhabiting regions with higher crop return have higher levels of agricultural intensity. In particular, a one-standard deviation increase in crop return is associated with a 0.24 standard deviations increase in agricultural intensity.²⁹ Moreover, columns (3)-(5) document the positive association between the

languages) (94, 95).

²⁹Table S7.5 documents the robustness of this result, by constraining the analysis to the set of societies that

Table S4.3: Geographic Origins of Agricultural Intensity and Periphrastic Future Tense

	Mechanism				
	Agricultural Intensity		Future Tense		
	OLS	OLS	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)
Crop Return (pre-1500CE)	0.22*** (0.08)	0.25*** (0.05)			
Agricultural Intensity			0.07** (0.03)	0.08* (0.05)	0.21** (0.11)
All Geographic Controls	No	Yes	No	Yes	Yes
Regional FE	No	Yes	No	Yes	Yes
First-stage F-statistic					11.84
Adjusted- R^2	0.04	0.64	0.02	0.17	0.13
Observations	1283	1283	263	263	263

Notes: This table suggests, based on OLS and IV regressions, the positive statistically and economically significant association between pre-1500 crop return and the existence of a periphrastic future tense in a language. The first two columns provide the results of the association between crop return and the intensity of agricultural, and columns (3)-(5) provide evidence on the association between agricultural intensity and the emergence of periphrastic future tense. Heteroskedasticity robust standard error estimates clustered at the genus level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

level of agricultural intensity in an ethnic group and the probability of existence of periphrastic future tense in the language it speaks. The results suggest that a one standard deviation increase in agricultural intensity is associated with an 8 percentage points increase in the probability of the existence of periphrastic future tense in the case of OLS. Table S4.4 provides further evidence for the mechanism underlying this relation stemming from agriculture.

S4.D Time Preference and Periphrastic Future Tense

This section explores the association between the level of long-term orientation of the community that speaks a certain language and the existence of periphrastic future tense in the language. In light of the proposed hypothesis that the periphrastic future tense is complementary to long-term oriented behavior, the prevalence of long-term orientation within a given society ought to be associated with the presence of periphrastic future tense in the language used among speak languages for which data on the existence of periphrastic future tense is available.

Table S4.4: Geographic Origins of Agricultural Societies and Periphrastic Future Tense

	Mechanism				
	Agricultural Society		Future Tense		
	OLS	OLS	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)
Crop Return (pre-1500CE)	0.10*** (0.04)	0.11*** (0.03)			
Agricultural Society			0.20* (0.12)	0.17 (0.14)	0.55** (0.27)
All Geographic Controls	No	Yes	No	Yes	Yes
Regional FE	No	Yes	No	Yes	Yes
First-stage F-statistic					19.78
Adjusted- R^2	0.06	0.55	0.01	0.17	0.14
Observations	1181	1181	242	242	242

Notes: This table suggests, based on OLS and IV regressions, the positive statistically and economically significant association between pre-1500 crop return and the existence of a periphrastic future tense in a language. The first two columns provide the results of the association between crop return and being an agricultural society, and columns (3)-(5) provide evidence on the association between being an agricultural society and the emergence of periphrastic future tense. Heteroskedasticity robust standard error estimates clustered at the genus level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

members of this society. Indeed, the common geographical origins (i.e., the natural return to agricultural investment) of long-term orientation as well as periphrastic future tense suggest that this positive association is likely to be observed across speech communities.

Table S4.5 explores the association between the prevalence of long-term orientation within a given society and the presence of periphrastic future tense in the language used among members of this society.³⁰ It documents the positive significant association between the existence of the periphrastic future tense and long-term orientation. This positive association is present among speakers of languages in the world as a whole (Columns (1) and (2)), as well as among speakers who reside in the Old World and speak languages originated in the Old World (Columns (3) and (4)). Moreover, as reported in Columns (2) and (4), this association is highly significant and stable in magnitude if one accounts for the geographical origins (i.e., the natural return to

³⁰The measure of Long-Term Orientation is based on the following question in the World Values Survey (WVS): “Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important?” Individuals are considered to have Long-Term Orientation if they answered “Thrift, saving money and things”.

Table S4.5: Long-Term Orientation and Periphrastic Future Tense

	Long-Term Orientation			
	All		Old World	
	(1)	(2)	(3)	(4)
Periphrastic Future Tense	0.12** (0.05)	0.15*** (0.04)	0.15*** (0.04)	0.17*** (0.05)
Crop Return (pre-1500CE)		0.29* (0.16)		0.26 (0.18)
Regional FE	Yes	Yes	Yes	Yes
Main Geographical Controls	No	Yes	No	Yes
Adjusted- R^2	0.16	0.24	0.18	0.26
Observations	106	106	100	100

Notes: This table documents the statistically and economically significant positive association between periphrastic future tense and long-term orientation. The results account for the geographical origin of the periphrastic future tense and other geographical characteristics of the contemporary homeland of the language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

agricultural investment) of these two traits. These findings lend credence to the proposed hypothesis that the prevalence of the trait of long-term orientation and the presence of periphrastic future tense have coevolved in the course of human history.

S5 The Origins of Sex-Based Grammatical Gender Systems

The hypothesized coevolution of gender biases and sex-based grammatical gender suggests that in societies characterized by distinct gender roles and consequently by the existence of gender biases in the distant past, sex-based grammatical gender has emerged and persisted over time, reinforcing existing gender biases in society. Thus, in view of the suggested positive association between the suitability of land for the usage of the plow and the emergence of distinct gender roles in society (3, 15), the suitability of the land for the usage of the plow is hypothesized to affect the emergence and the prevalence of sex-based grammatical gender in the course of

human history.^{31,32,33}

This section analyzes the relation between the existence of sex-based grammatical gender across contemporary languages and the potential historical suitability of the land for the usage of the plow in their contemporary geographical homelands (i.e., the indigenous regions that spoke the language in the pre-colonial as well as in the post-colonial era), as well as in their ancestral homeland (i.e., their *Urheimat*). As explained in section S1.B.2, since plow positive crops are characterized by limited regional variation (e.g., they are entirely absent from the Americas in the pre-1500 period), the analysis will focus on the adverse effect of the average suitability of plow negative crops on the adoption of the plow, as well as the positive effect of the average suitability of all crops on the desirability of the adoption of agricultural technology such as the plow.³⁴

Table S5.1 explores the association between the potential historical suitability of the land for the usage of the plow and the existence of sex-based grammatical gender.³⁵ As suggested by column (1), the association between the potential caloric yield of plow negative crops in the pre-1500CE period in the language's contemporary homeland and the existence of sex-based grammatical gender is negative and statistically significant at the 5% level, suggesting that a one standard deviation increase in the potential caloric yield of plow negative crops in the language's contemporary homeland is associated with a 13 percentage points decrease in the probability that the language is characterized by a sex-based grammatical gender. Moreover, the average caloric suitability of all crops in the pre-1500CE period in the language's contemporary homeland is indeed positively and statistically significantly associated with the existence of sex-based

³¹Differences in labor productivity across genders may emerge due to a variety of sources in agricultural and non-agricultural societies. However, in light of the importance attributed to the plow in the emergence of a gender gap in labor productivity in agriculture (15), and in view of data availability, the analysis centers around those geographical characteristics that were conducive for the use of the plow. The presence of these additional forces could have further enriched the analysis and provide additional documentation to the geographical origins of the coevolution of cultural and linguistic characteristics. However, the lack of extensive data on exogenous geographical sources of variation in differential productivity across gender for non-agricultural societies prevents such an analysis.

³²Although the analysis focuses on the plow, animal traction also may have played an important role for understanding gender biases (48).

³³As an example, consider two Asian languages: Hebrew and Western Farsi. Hebrew is spoken in Israel, where the land was suitable for the use of the plow and plow was historically present. According to our theory, the Hebrew language is predicted to have sex-based grammatical gender, which is indeed the case. Western Farsi is spoken in Iran; where the land was not suitable for the use of the plow and plow was not used historically. According to our theory, Western Farsi is predicted not to have sex-based grammatical gender, which is indeed the case.

³⁴In the main body of the paper the analysis reverse coded the potential caloric plow suitability in order to simplify exposition. The analysis in the Supplemental Material maintains the original coding.

³⁵Table S8.1 reports the estimated coefficients on the control variables in Table S5.1.

grammatical gender, suggesting that a one standard deviation increase in the potential caloric yield of all crops in the language's contemporary homeland is associated with a 17 percentage points increase in the probability that the language has sex-based grammatical gender. Furthermore, as reported in column (2), the estimated relationship becomes more significant once the potentially confounding effects of the geographical characteristics (i.e., absolute latitude, mean elevation, terrain ruggedness, and the length of its sea coast) in the contemporary homeland of each language, are accounted for.

Table S5.1: Geographic Origins of Sex-Based Grammatical Gender (OLS)

	Existence of Sex-Based Grammatical Gender					
	(1)	(2)	(3)	(4)	(5)	(6)
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.13** (0.06)	-0.19*** (0.06)	-0.25*** (0.06)	-0.26*** (0.06)	-0.29*** (0.06)	-0.23*** (0.08)
All Crops (Average Caloric Yield, pre-1500)	0.17*** (0.06)	0.21*** (0.06)	0.28*** (0.06)	0.29*** (0.06)	0.32*** (0.06)	0.25*** (0.08)
Geographical Controls	No	Yes	Yes	Yes	Yes	Yes
Precipitation Controls	No	No	Yes	Yes	Yes	Yes
Temperature Controls	No	No	No	Yes	Yes	Yes
Unproductive Period	No	No	No	No	Yes	Yes
Regional FE	No	No	No	No	No	Yes
Adjusted- R^2	0.03	0.06	0.14	0.15	0.15	0.21
Observations	217	217	217	217	217	217

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of sex-based grammatical gender in a language, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S4.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a sex-based grammatical gender in the language. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Since the potential historical agricultural determinants of plow usage are estimated based on agro-climatic conditions, they might be capturing the potentially direct impact of climatic conditions on the existence of sex-based grammatical gender. In particular, variability of climatic conditions may affect the riskiness of agricultural investment, and therefore the effective caloric yield of plow negative crops, while their spatial correlation may affect the potential for spatial diversification of risk. Nevertheless, as reported in columns (3) and (4), accounting for the potentially confounding effects of average temperature and precipitation, as well as their standard deviations, volatility, and spatial correlation, does not alter the qualitative results.

Furthermore, the length of the period within each year in which the land is not suitable for agriculture – the unproductive period – may have conceivably affected the desirability of the usage of the plow. Nevertheless, as indicated by column (5), accounting for the length of the unproductive period, reinforces the findings. Finally, as reported in column (6), accounting for regional fixed-effects, and therefore for unobserved time-invariant heterogeneity at the regional level does not alter the qualitative results. In particular, a one standard deviation increase in the pre-1500CE potential caloric yield of plow negative crops in the language’s contemporary homeland is associated with a 23 percentage points decrease in the probability that the language is characterized by a sex-based grammatical gender. Similarly, a one standard deviation increase in the pre-1500CE potential caloric yield of all crops in the language’s contemporary homeland is associated with a 25 percentage points increase in the probability that the language is characterized by a sex-based grammatical gender.

S5.A Robustness

The findings in Table S5.1 lend credence to the hypothesis that the potential historical agricultural determinants of plow usage in the language’s contemporary homeland, and their effect on gender biases, contributed to the emergence of sex-based grammatical gender. Nevertheless, the results conceivably might be biased due to omitted institutional, cultural, and human characteristics, precluding a causal interpretation of the estimated coefficients. While concerns about omitted variables will be directly addressed and mitigated in the next subsection, the remaining analysis in this section documents the robustness of the results to: (i) an alternative estimation method, (ii) spatial autocorrelation, and (iii) clustering of standard errors, while demonstrating that the findings are unlikely to be driven by selection on unobservables. Moreover, the robustness to alternative aspects of grammatical gender are also considered.

S5.A.1 Estimation Method, Spatial-Autocorrelation and Selection on Unobservables

Table S8.2 (Panel A) documents that the estimated average marginal effects of the Probit model are nearly identical to the estimated coefficients using OLS. Furthermore, Table S8.2 (Panel B) documents the robustness of the results to the clustering of the standard errors at the level of the language genus (i.e., a group of phylogenetically related languages inside a linguistic family), as well as spatial auto-correlation. In particular, the statistical significance is unaffected by the various methods used.

Table S8.2 (Panel B) further explores the potential bias that could be generated by omitted variables. In particular, using statistics on selection on observables and unobservables (91,92), it suggests that the degree of omitted variable bias is low and is unlikely to explain the magnitude of the estimated association between the potential historical agricultural determinants of plow usage and the existence of sex-based grammatical gender. In particular, omitted factors would need to be 1.3-2.3 times more strongly and negatively correlated with the caloric yield of plow negative crops, in order to account for the estimated association between caloric yield of plow negative crops and the emergence of sex-based grammatical gender. Similarly, omitted factors would need to be 1.6-3 times more strongly and negatively correlated with the caloric yield of all crops, in order to account for the estimated association between caloric yield of all crops and the emergence of sex-based grammatical gender. Thus, the estimated coefficients could be considered as the lower bounds in absolute value of the true association. The bias-adjusted estimated associations between potential historical agricultural determinants of plow usage and the existence of sex-based grammatical gender are larger in absolute value than the OLS estimate. In particular, as reported in column (6), following (92), the bias-corrected association between caloric yield of plow negative crops and the existence of sex-based grammatical gender, assuming that the unobservables are as strongly correlated with pre-1500CE caloric yield of plow negative crops as the set of observables that are accounted for, implies that a one standard deviation increase in caloric yield of plow negative crops in the language's contemporary homeland is associated with a 27 percentage points decrease in the probability of the existence of sex-based grammatical gender in the language.

S5.A.2 Alternative Aspects of Grammatical Gender

Tables S8.4, S8.8 and S8.12 document qualitatively similar associations between the geographical determinants of plow-based agriculture and (i) an alternative measure of sex-based grammatical gender, (ii) the intensity of grammatical gender in a language, and (iii) the existence of gender distinctions in personal pronouns, respectively. These results suggest that various aspects of grammatical gender that have reinforced gender biases, were associated with the geographical determinants of plow-based agriculture. Moreover, Tables S8.2, S8.5, S8.9 and S8.13 establish the robustness of these alternative associations to: (i) an alternative estimation method, (ii) spatial autocorrelation, and (iii) clustering of standard errors, while demonstrating that the findings are unlikely to be driven by selection on unobservables.

S5.A.3 Agricultural Determinants of Plow Usage and Other Language Structures

The associations between the potential historical agricultural determinants of plow usage and the existence of sex-based grammatical gender may reflect a broader association between plow-based agriculture and a wide range of linguistic traits, rather than its unique association with sex-based grammatical gender, weakening the proposed hypothesis. Moreover, the association between the suitability of land for the usage of the plow in the language's contemporary homeland and the existence of sex-based grammatical gender in the language may reflect a broader association between a whole range of linguistic traits and gender biases.

Table S14.3 suggests, however, that alternative language structures (as reported by WALS), such as the existence of other temporal and non-temporal structures, are largely uncorrelated with the geographical determinants of plow-based agriculture. In particular the potential historical agricultural determinants of plow usage is not significantly associated with the existence of the perfect tense, and it is orthogonal to non-temporal linguistic characteristics such as: (i) the existence of possessive classifications, (ii) the existence of coding for evidentiality, (iii) the number of consonants, and (iv) the number of colors. Furthermore, using all available data on language structures as reported in WALS, Figure S14.4 documents the orthogonality of the potential historical agricultural determinants of plow usage for nearly all language structures except for sex-based grammatical gender.³⁶ These results lend further support to the hypothesis that the potential historical suitability of the land for the usage of the plow is associated with the emergence of the sex-based grammatical gender system.

S5.B Geographical Origins: Contemporary Homeland vs. Urheimat

This subsection exploits the descent of contemporary languages from proto-languages to explore the relative contributions of the potential suitability of land for the usage of the plow in (i) the contemporary homeland of the language and (ii) the ancestral homeland of the language (i.e., its Urheimat) to the existence of a sex-based grammatical gender in daughter languages.

This analysis mitigates concerns about the role of omitted variables in the associations between the potential historical suitability of land for the usage of the plow and the existence of sex-based grammatical gender, and the sorting of individuals into geographical niches that com-

³⁶Specifically, the only other structures that are statistically associated with the potential historical agricultural determinants of plow usage are: perfective and imperfective aspect (variable 65A in WALS) and existence of past tense (66A).

plement their linguistic and cultural traits. Moreover, it sheds light on the historical depth of the observed relationship between the potential suitability of land for the usage of the plow and the existence of sex-based grammatical gender, and the role of the culturally-embodied effects of geographical characteristics, rather than the geography per se, in the evolution of sex-based grammatical gender.

The analysis mirrors the epidemiological approach to cultural diffusion, and suggests the geographical origins of the existence of sex-based grammatical gender based on variation across languages located outside the Urheimat (i.e., ancestral homeland of their proto-language), permitting the analysis to account for host region fixed-effects and thus to capture unobserved time-invariant heterogeneity at the regional level, thus mitigating concerns about omitted variables.

Table S5.2 explores the relative contributions of the pre-1500CE potential suitability of land for the usage of the plow in the contemporary language's homeland vs. the Urheimat to the presence of sex-based grammatical gender in daughter languages. In particular, it documents that the existence of sex-based grammatical gender among daughter languages located outside their Urheimat is associated predominantly with the potential suitability of land for the usage of the plow in the Urheimat.³⁷

The analysis is conducted in two layers, distinguished by the geographical proximity of the daughter languages to their Urheimat. Columns (1)-(3) consider daughter languages anywhere outside of the Urheimat, while columns (4)-(6) consider languages at a migratory distance of one week from the Urheimat (i.e, approximately 1300km away from the Urheimat). The analysis suggests that while the association between potential suitability of land for the usage of the plow in the contemporary homeland of a language and the existence of sex-based grammatical gender among migrant languages is fragile (columns (1) and (4)), the association between the Urheimat's crop return and the existence of sex-based grammatical gender among daughter languages located outside of the Urheimat is larger and highly significant (column (2) and (5)). Moreover, the impact of the potential suitability of land for the usage of the plow in the Urheimat remains significant, while the impact of crop return in the contemporary homeland remains fragile if the geographical characteristics of the contemporary homeland of the language and its Urheimat are considered jointly.³⁸

³⁷Given the potential dependence of languages within the same language family, standard errors are clustered at this level.

³⁸Given the greater uncertainty about the geographical extent of the Urheimat, as opposed to the contemporary linguistic homeland, measurement errors in the estimates of crop return in the Urheimat are likely to be larger than those from crop return in the contemporary homeland and a-priori the horse race between the two variables should

Table S5.2: The Persistent Impact of Geographical Characteristics in the Urheimat, rather than in the Contemporary Linguistic Homeland, on the Existence of Sex-Based Grammatical Gender:

Daughter Languages Outside the Urheimat

	The Existence of Sex-Based Gender System					
	Migratory Distance to the Urheimat					
	Any Distance			At Least 1 Week		
	Homeland	Urheimat	Both	Homeland	Urheimat	Both
	(1)	(2)	(3)	(4)	(5)	(6)
Homeland's Plow Negative Crops (Avg. Caloric Yield, pre-1500)	-0.24*** (0.07)		-0.11 (0.07)	-0.12 (0.09)		-0.13** (0.05)
Homeland's Crops (Average Caloric Yield, pre-1500)	0.27*** (0.07)		0.15** (0.07)	0.11 (0.09)		0.09* (0.05)
Urheimat's Plow Negative Crops (Avg. Caloric Yield, pre-1500)		-0.35*** (0.09)	-0.31*** (0.10)		-0.51*** (0.18)	-0.47** (0.21)
Urheimat's Crops (Average Caloric Yield, pre-1500)		0.53*** (0.10)	0.50*** (0.11)		0.64*** (0.20)	0.60** (0.23)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes
Homeland Geographical Characteristics	Yes	No	Yes	Yes	No	Yes
Urheimat Geographical Characteristics	No	Yes	Yes	No	Yes	Yes
Adjusted- R^2	0.23	0.42	0.41	0.30	0.53	0.50
Observations	195	195	195	131	131	131
Language Families	57	57	57	29	29	29

Notes: This table documents the importance of agricultural productivity and the suitability of the land for the usage of the plow in the Urheimat, rather than the contemporary homeland, to the existence of sex-based grammatical gender in daughter languages located outside of the Urheimat, accounting for regional fixed-effects and other geographical characteristics. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length. Columns (1)-(3) consider daughter languages anywhere outside of the Urheimat while column (4)-(6) those who are at least a week of migratory distance from the Urheimat. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

The analysis suggests that the existence of sex-based grammatical gender began its formation in the distant past when the speakers of the proto-language still inhabited the Urheimat of the language, reflecting the portable culturally-embodied effect of the suitability of the land for the usage of the plow rather than the direct effect of geography.³⁹ As discussed in Appendix S12, the timing of the emergence of the plow is consistent with the subsequent emergence of

have resulted in fact in the decline in the significance of crop return in the Urheimat.

³⁹Tables S8.7, S8.10 and S8.15 examine the robustness of this analysis to alternative aspects of grammatical gender.

sex-based grammatical gender in a significant number of languages. The findings suggest that the results are driven by neither omitted factors nor geographical sorting, and it provides support for the deep-historical origins of the structure of sex-based grammatical gender.

S5.C Mechanism

This section presents additional supportive evidence for the hypothesized geographical origins of the sex-based grammatical gender. In view of the evidence about the effect of plow usage on gender biases (3, 15), the effect of the geographical determinants of plow-based agriculture on the emergence of sex-based grammatical gender would necessitate the adoption and use of the plow. In particular, as illustrated in Figure S5.1, the proposed hypothesis would imply that: (i) the geographical determinants of plow-based agriculture ought to be associated with the adoption and use of the plow, and (ii) the use of the plow ought to be associated with the existence of sex-based grammatical gender.

Figure S5.1: Geographical Determinants of Plow-based Agriculture, Plow Usage & Grammatical Gender

Table S5.3 presents supporting evidence for the proposed mechanism. In particular, it documents a significant association between the geographical determinants of plow-based agriculture and the usage of the plow in pre-colonial societies, which in turn is positively associated with the emergence of sex-based grammatical gender. Using data from the Ethnographic Atlas (51) on the presence and use of aboriginal plows, the analysis explores the association between the geographical determinants of plow-based agriculture and a pre-colonial society's plow usage. In particular, columns (1) and (2) suggest that societies that resided in regions that were more suitable for the use of the plow are more likely to use an aboriginal plow. A one-standard deviation increase in pre-1500CE caloric suitability of plow negative crops is associated with a 6 percentage points decrease in the probability of the usage of the plow. Similarly,

Table S5.3: Geographic Origins of Plow Usage and Sex-Based Grammatical Gender

	Mechanism				
	Plow		Grammatical Gender		
	OLS	OLS	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.26*** (0.04)	-0.06* (0.03)			
Average Caloric Yield (All Crops, pre-1500)	0.26*** (0.04)	0.10*** (0.03)			
Aboriginal Plow			0.38*** (0.08)	0.23* (0.13)	0.95*** (0.35)
All Geographic Controls	No	Yes	No	Yes	Yes
Regional FE	No	Yes	No	Yes	Yes
First-stage F-statistic					15.45
Adjusted- R^2	0.19	0.47	0.14	0.31	0.06
Observations	1158	1158	145	145	145

Notes: This table documents, based on OLS regression, the positive statistically and economically significant association between the geographical determinants of plow usage and the existence of sex-based grammatical gender in a language. The first two columns provide the associations between the geographical determinants of plow suitability and actual usage of the plow, and columns (3)-(5) provide evidence on the association between actual usage of the plow and the emergence of sex-based grammatical gender. Heteroskedasticity robust standard error estimates clustered at the genus level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

a one-standard deviation increase in pre-1500CE caloric suitability of all crops is associated with a 9 percentage points increase in the probability of the usage of the plow. Moreover, columns (3)-(5) suggests that there exist a positive association between the usage of the plow in a pre-colonial ethnic group and the probability of the existence of sex-based grammatical gender in the language used by this ethnic group. The results suggest that a one standard deviation increase in the usage of the plow is associated with a 23 percentage points increase in the probability of the existence of sex-based grammatical gender in the case of OLS.

S5.D Gender Bias and Sex-Based Grammatical Gender

This section explores the association between the existence of a gender bias in a given community and the existence of sex-based grammatical gender in the language the members of the community speak. In light of the proposed hypothesis that sex-based grammatical gender is complementary to gender biased behavior, the prevalence of gender biases within a given society is likely to be associated with the presence of sex-based grammatical gender in the lan-

guage used among members of this society. Indeed, the common geographical origins (i.e., the geographical determinants of plow-based agriculture) of gender bias as well as sex-based grammatical gender suggest that this positive association is likely to be observed across various speech communities.

Table S5.4: Gender Bias and Sex-based Grammatical Gender

	Gender Bias			
	All		Old World	
	(1)	(2)	(3)	(4)
Existence of Sex-Based Grammatical Gender	0.27 (0.23)	0.42** (0.16)	0.18 (0.23)	0.18* (0.09)
Plow Negative Crops (Average Caloric Yield, pre-1500)		0.16 (0.43)		-0.25 (0.29)
All Crops (Average Caloric Yield, pre-1500)		-0.35 (0.49)		0.19 (0.34)
Main Geographical Controls	No	Yes	No	Yes
Adjusted- R^2	0.04	0.33	-0.00	0.38
Observations	33	33	30	30

Notes: This table documents the statistically and economically significant association between the existence of sex-based grammatical gender and gender bias in society. The analysis accounts for the geographical origins of sex-based grammatical gender and other geographical characteristics of the contemporary homeland of the language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S5.4 explores the association between the prevalence of gender biases within a given society and the presence of sex-based grammatical gender in the language used among members of this society.⁴⁰ It documents a positive association between the existence of sex-based grammatical gender and gender biases.⁴¹ This positive association is present among speakers of languages in the world as a whole (Columns (1) and (2)), as well as among speakers who reside in the Old World and speak languages originated in the Old World (Columns (3) and (4)). These

⁴⁰The measure of gender bias is based on the first principal component to the answer of the following questions in the World Values Survey (WVS): (i) “When jobs are scarce men should have more right to a job than women”, (ii) “Men make better political leaders than women do”, (iii) “Women should have the same rights as men”, (iv) “A job is the best way for women to be independent”, (v) “It is a problem if women have more income than husband”, and (vi) “Men make better political leaders than women do”.

⁴¹The sample of languages in the WVS for which grammatical gender data exists is small and has little regional variation. This impedes the use of regional fixed-effects in the analysis and reduces the significance of the estimates.

findings lend credence to the proposed hypothesis that the prevalence of gender biases and the presence of sex-based grammatical gender have coevolved in the course of human history.

S6 The Origins of Politeness Distinctions in Pronouns

The hypothesized coevolution of hierarchical orientation (e.g., obedience, conformity, and power distance) and politeness distinctions suggests that in societies characterized by higher hierarchical orientation in the distant past, politeness distinctions have emerged and persisted over time, reinforcing the existing hierarchical orientation in society. Thus, in view of the suggested positive association between the extent of ecological diversity and the emergence of trade and economic specialization, which in turn were conducive to the emergence of hierarchical societies (18, 97), ecological diversity is hypothesized to affect the coevolution of the prevalence of politeness distinctions and hierarchical orientation in the course of human history across the entire globe.⁴²

Table S6.1: Geographic Origins of Politeness Distinctions (OLS)

	Existence of Politeness Distinctions					
	(1)	(2)	(3)	(4)	(5)	(6)
Ecological Diversity	0.15*** (0.03)	0.10*** (0.04)	0.10*** (0.04)	0.10** (0.04)	0.10** (0.04)	0.10** (0.04)
Geographical Controls	No	Yes	Yes	Yes	Yes	Yes
Precipitation Controls	No	No	Yes	Yes	Yes	Yes
Temperature Controls	No	No	No	Yes	Yes	Yes
Unproductive Period	No	No	No	No	Yes	Yes
Regional FE	No	No	No	No	No	Yes
Adjusted- R^2	0.09	0.18	0.20	0.20	0.21	0.32
Observations	198	198	198	198	198	198

Notes: This table documents the positive, statistically, and economically significant association between ecological diversity and the existence of politeness distinctions in a language, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S4.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

⁴²Although the analysis focuses on the ecological diversity, animal traction also may have played an important role for understanding the emergence of social inequality (48).

This section analyzes the relation between the existence of politeness distinctions in pronouns across contemporary languages and the extent of ecological diversity in their contemporary geographical homelands (i.e., the indigenous regions that spoke the language in the pre-colonial as well as in the post-colonial era), as well as in their ancestral homeland (i.e., their *Urheimat*). In particular, Table S6.1 explores the association between the extent of ecological diversity in the contemporary homeland and the existence of politeness distinctions.⁴³ As suggested by column (1), the association between the extent of ecological diversity in the language's contemporary homeland and the existence of politeness distinctions is positive and statistically significant at the 1% level, suggesting that a one standard deviation increase in the extent of ecological diversity in the language's contemporary homeland is associated with a 15 percentage points increase in the probability that the language is characterized by the existence of politeness distinctions. Furthermore, as reported in column (2), the estimated relationship remains highly significant once the potentially confounding effects of the geographical characteristics (i.e., absolute latitude, mean elevation, terrain ruggedness, and the length of its sea coast) in the contemporary homeland of each language, are accounted for. Moreover, since ecological diversity partly captures climatic conditions, it might be capturing the potentially direct impact of climatic conditions on the existence of politeness distinctions. Nevertheless, as reported in columns (3) and (4), accounting for the potentially confounding effects of average temperature and precipitation, as well as their standard deviations, volatility, and spatial correlation, does not alter the results qualitatively or quantitatively.

Furthermore, the length of the period within each year in which the land is not suitable for agriculture – the unproductive period – may have conceivably affected the profitability of trade and therefore the association between ecological diversity and the emergence of hierarchical societies. Nevertheless, as indicated by column (5), accounting for the length of the unproductive period has little impact on the estimated association. Finally, as reported in column (6), accounting for regional fixed-effects, and therefore for unobserved time-invariant heterogeneity at the regional level does not alter the qualitative results. In particular, a one standard deviation increase in the extent of ecological diversity in the language's contemporary homeland is associated with a 10 percentage points increase in the probability that the language is characterized by politeness distinctions.

⁴³Table S9.1 reports the estimated coefficients on the control variables in Table S6.1.

S6.A Robustness

The findings in Table S6.1 lend credence to the hypothesis that the extent of ecological diversity and its effect on the emergence of hierarchical societies, contributed to the emergence of politeness distinctions. Nevertheless, the results conceivably might be biased due to omitted institutional, cultural, and human characteristics, precluding a causal interpretation of the estimated coefficients. While concerns about omitted variables will be directly addressed and mitigated in the next subsection, the remaining analysis in this section documents the robustness of the results to: (i) an alternative estimation method, (ii) spatial autocorrelation, and (iii) clustering of standard errors, while demonstrating that the findings are unlikely to be driven by selection on unobservables. Moreover, the robustness to alternative aspects of politeness distinctions are also considered.

S6.A.1 Estimation Method, Spatial-Autocorrelation and Selection on Unobservables

Table S9.2 (Panel A) documents that the estimated average marginal effects of the Probit model are nearly identical to the estimated coefficients using OLS. Furthermore, Table S9.2 (Panel B) documents the robustness of the results to the clustering of the standard errors at the level of the language genus (i.e., a group of phylogenetically related languages inside a linguistic family), as well as spatial auto-correlation. In particular, the statistical significance is unaffected by the various methods used.

Table S9.2 (Panel B) further explores the potential bias that could be generated by omitted variables. In particular, using statistics on selection on observables and unobservables (91, 92), it documents that the degree of omitted variable bias is low and is unlikely to explain the magnitude of the estimated association between ecological diversity and the existence of politeness distinctions. In particular, omitted factors would need to be 3-5 times more strongly correlated with ecological diversity, in order to account for the estimated association between ecological diversity and the emergence of politeness distinctions. Indeed, the bias-adjusted estimated coefficient of ecological diversity is only slightly smaller than the value of the OLS estimate. In particular, as reported in column (6), following (92), the bias-corrected coefficient of ecological diversity, assuming that the unobservables are as strongly correlated with ecological diversity as the set of observables that are accounted for, implies that a one standard deviation increase in ecological diversity is associated with an 8 percentage points increase in the probability of the existence of politeness distinctions in pronouns.

S6.A.2 Alternative Aspects of Politeness Distinctions

Tables S9.4 and S9.10 document qualitatively similar associations between ecological diversity and (i) an alternative measure of politeness distinctions in pronouns, and (ii) the intensity of politeness distinctions in a language. These results suggest that various aspects of politeness distinctions that have reinforced hierarchical orientation, were associated with ecological diversity. Moreover, Tables S9.5 and S9.11 establish the robustness of these alternative associations to: (i) an alternative estimation method, (ii) spatial autocorrelation, and (iii) clustering of standard errors, while demonstrating that the findings are unlikely to be driven by selection on unobservables.

S6.A.3 Ecological Diversity and Other Language Structures

The associations between the extent of ecological diversity and politeness distinctions may reflect a broader association between ecological diversity and a wide range of linguistic traits, rather than its unique association with politeness distinctions, weakening the proposed hypothesis. Moreover, the association between the extent of ecological diversity and politeness distinctions in the language may reflect a broader association between a whole range of linguistic traits and politeness distinctions.

Table S14.7 suggest, however, that alternative language structures (as reported by WALs), such as the existence of other temporal and non-temporal structures, are largely uncorrelated with the geographical determinants of hierarchical society. In particular the extent of ecological diversity is not significantly associated with the existence of the past and the perfect tense, and it is orthogonal to non-temporal linguistic characteristics such as: (i) the existence of possessive classifications, (ii) the existence of coding for evidentiality, (iii) the number of consonants, and (iv) the number of colors. Furthermore, using all available data on language structures as reported in WALs, Figure S14.5 documents the orthogonality of ecological diversity for all language structures that are not related to politeness distinctions.⁴⁴ These results lend further

⁴⁴Specifically, the only other structures that are statistically associated with ecological diversity are: consonant-vowel ratio (variable 3A in WALs), inclusive/exclusive distinction in independent personal pronouns (39A), and existence of passive constructions (107A). The significant correlation with the consonant-vowel ratio may reflect the positive impact of this variable on the emergence of states (97), and thus of hierarchical societies, which in turn are also associated with politeness distinctions. The association with inclusive/exclusive distinction in independent personal pronouns (39A) may reflect the commonality among certain aspects of languages or simply the fact that this variable is not ordinal, and as such the correlation is merely due to randomness in the ordering of this variable. Finally, the association with the existence of passive constructions seems to reflect its use as a politeness device (5, 99).

support to the hypothesis that ecological diversity is uniquely associated with the emergence of politeness distinctions.

S6.B Geographical Origins: Contemporary Homeland vs. Urheimat

This subsection exploits the descent of contemporary languages from proto-languages to explore the relative contributions of ecological diversity in (i) the contemporary homeland of the language and (ii) the ancestral homeland of the language (i.e., its Urheimat) to the existence of politeness distinctions.

Table S6.2: The Relative Contribution of Ecological Diversity in the Linguistic Homeland vs. the Urheimat to the Existence of Politeness Distinction:
Daughter Languages Outside Urheimat

	The Existence of Politeness Distinctions					
	Migratory Distance to Urheimat					
	Any Distance			At Least 1 Week		
	Homeland	Urheimat	Both	Homeland	Urheimat	Both
	(1)	(2)	(3)	(4)	(5)	(6)
Homeland's Ecological Diversity	0.10*** (0.03)		0.07* (0.04)	0.12*** (0.02)		0.10*** (0.02)
Urheimat's Ecological Diversity		0.09*** (0.03)	0.05 (0.03)		0.16*** (0.04)	0.07 (0.05)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes
Homeland Geographical Characteristics	Yes	No	Yes	Yes	No	Yes
Urheimat Geographical Characteristics	No	Yes	Yes	No	Yes	Yes
Adjusted- R^2	0.34	0.37	0.41	0.37	0.43	0.49
Observations	180	180	180	126	126	126
Language Families	54	54	54	28	28	28

Notes: This table documents the importance of ecological diversity in the contemporary homeland rather than in the Urheimat to the presence of politeness distinctions in daughter languages located outside of the Urheimat, accounting for regional fixed-effects and other geographical characteristics. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length. Columns (1)-(3) consider daughter languages anywhere outside of the Urheimat while column (4)-(6) those who are at least a week of migratory distance from the Urheimat. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S6.2 explores the relative contributions of the extent of ecological diversity in the contemporary language's homeland vs. its Urheimat to the presence of politeness distinctions in

daughter languages. It suggests that the existence of politeness distinctions in pronouns among daughter languages located outside their Urheimat are predominantly associated with ecological diversity in their contemporary homeland.⁴⁵ Columns (1) - (6) suggest that the extent of ecological diversity in the contemporary homeland of a language has a highly significant positive association with the existence of politeness distinctions among migrant languages, dominating the association with the characteristics in the Urheimat. Accounting for regional fixed-effects and other geographical characteristics of the homeland, a one standard deviation increase in ecological diversity in the contemporary homeland of the language is associated with a 10-12 percentage points increase in the probability that a migrant daughter language would have politeness distinctions in pronouns. Similarly, Tables S9.7 and S9.12 document analogous results of the association between ecological diversity and (i) an alternative measure of politeness distinctions in pronouns and (ii) the intensity of politeness distinctions in pronouns.

The analysis mirrors the epidemiological approach to cultural diffusion, and suggests the geographical origins of the existence of politeness distinctions based on variation across languages located outside the Urheimat (i.e., ancestral homeland of their proto-language). This approach permits the analysis to account for host region fixed-effects and thus to capture unobserved time-invariant heterogeneity at the regional level, thus mitigating concerns about omitted variables.

These results suggests that politeness distinctions have emerged primarily during the demic diffusion of the language. Indeed, these results can be interpreted as an outcome of sorting of hierarchy-averse individuals into geographical locations where hierarchies are less likely to emerge. Interestingly, this would suggest that hierarchical structures and thus politeness distinctions would be more prevalent in geographical locations characterized by ecological diversity, where individuals are not freely mobile and could not flee oppression (100).

The evidence provided in this section supports the proposed hypothesis that ecological diversity affects the existence of politeness distinctions in a language, shedding light on the co-evolution of language and culture. Furthermore, consistent with evidence about the greater adaptability of politeness distinctions to environmental and political changes (30–32, 41, 42), in comparison to the structure of the future tense and sex-based grammatical gender, the findings suggest that while geographical characteristics in the Urheimat are associated with the existence of the periphrastic future tense and sex-based grammatical gender, the geographical character-

⁴⁵Given the potential dependence of languages within the same language family, standard errors are clustered at this level.

istics in the homeland of the daughter languages are associated with the existence of politeness distinctions. In particular, in line with the conventional wisdom among linguists about the relative fluidity of politeness distinctions, the evidence suggests that the existence of politeness distinctions in contemporary languages are originated predominantly in cultural processes that took place after proto-languages were formed and daughter languages had started to emerge. Moreover, the evidence is consistent with the possible sorting of hierarchy-averse individuals into geographical locations that are less prone to be conducive for the emergence of hierarchical structures, which may have played a role in the positive association between ecological diversity in the contemporary homeland of the language and the existence of politeness distinctions.

S6.C Mechanism

This section presents additional supportive evidence for the hypothesized geographical origins of politeness distinctions. In view of the evidence about the effect of ecological diversity on the emergence of hierarchical structures (18, 97), the effect of ecological diversity on the emergence of politeness distinctions would necessitate the emergence of hierarchical structures. In particular, as illustrated in Figure S6.1, the proposed hypothesis would imply that: (i) ecological diversity ought to be associated with more hierarchical structures, and (ii) more hierarchical structures ought to be associated with the existence of politeness distinctions.

Figure S6.1: Ecological Diversity, Hierarchy & Politeness Distinctions

Table S6.3 presents supporting evidence for the proposed mechanism. In particular, it documents the robust positive association between ecological diversity and the level of jurisdictional hierarchy in a pre-colonial ethnic group, which in turn is positively associated with the emergence of politeness distinctions in pronouns. Using data from the Ethnographic Atlas (51) on the level of jurisdictional hierarchy above the local level – a proxy for the size and strength of the State and hierarchical structures – the analysis explores the association between ecological

diversity and a pre-colonial ethnicity's level of jurisdictional hierarchy. In particular, columns (1) and (2) suggest that societies inhabiting regions with higher ecological diversity had higher levels of jurisdictional hierarchy. The results suggest that a one-standard deviation increase in ecological diversity is associated with 0.12 standard deviations increase in jurisdictional hierarchy. Moreover, columns (3) and (4) document a positive association between the level of jurisdictional hierarchy in a pre-colonial ethnic group and the probability of existence of politeness distinctions in the language spoken by the ethnic group. The results suggest that a one standard deviation increase in jurisdictional hierarchy is associated with a 16 percentage points increase in the probability of existence of politeness distinctions in pronouns.

Table S6.3: Geographic Origins of Politeness Distinctions and Jurisdictional Hierarchy

	Mechanism				
	Jurisdictional Hierarchy		Politeness Distinctions		
	OLS	OLS	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)
Ecological Diversity	0.24*** (0.06)	0.12*** (0.04)			
Jurisdictional Hierarchy			0.21*** (0.02)	0.16*** (0.03)	0.32*** (0.06)
All Geographic Controls	No	Yes	No	Yes	Yes
Regional FE	No	Yes	No	Yes	Yes
First-stage F-statistic					14.24
Adjusted- R^2	0.07	0.32	0.34	0.48	0.35
Observations	1137	1137	139	139	139

Notes: This table documents the positive statistically and economically significant association between the geographical determinants of statehood, as measured by jurisdictional hierarchy beyond the local level, and politeness distinctions in a language. The first two columns provide the results of the association between ecological diversity and statehood, and columns (3) and (4) show the association between statehood and the emergence of politeness distinctions. Column (5) tests the aforementioned mechanism using an IV strategy. The table shows the estimated coefficients in an OLS regression as the dependent variable in columns (1) and (2) is not binary. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

S6.D Hierarchical Orientation and Politeness Distinctions

This section explores the association between the hierarchical orientation in a given community and the existence of politeness distinctions in the language the members of the community speak. In light of the proposed hypothesis that politeness distinctions in pronouns are complementary to hierarchical structures and cultures, the prevalence of hierarchical orientation within

a given society is likely to be associated with the presence of politeness distinctions in the language used among members of this society. Indeed, the common geographical origins (i.e., ecological diversity) of hierarchical orientation as well as politeness distinctions suggest that this positive association is likely to be observed across various speech communities.

Table S6.4: Hierarchy and Politeness

	Hierarchy Index			
	All		Old World	
	(1)	(2)	(3)	(4)
Politeness Distinctions	0.37** (0.13)	0.23** (0.09)	0.45*** (0.10)	0.29*** (0.06)
Ecological Diversity		0.14 (0.13)		0.05 (0.14)
Main Geographical Controls	No	Yes	No	Yes
Adjusted- R^2	0.12	0.27	0.19	0.36
Observations	53	53	50	50

Notes: This table documents the statistically and economically positive association between politeness distinctions and preferences for hierarchy in society. The analysis accounts for the geographical origins of politeness distinctions and other geographical characteristics of the contemporary homeland of the language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S6.4 explores the association between the prevalence of hierarchical orientation within a given society and the presence of politeness distinctions in pronouns in the language used among members of this society.⁴⁶ It documents the positive statistically and economically significant association between the existence of the politeness distinctions in pronouns and hierarchical orientation.⁴⁷ This significant positive association is present among speakers of languages in the world as a whole (Columns (1) and (2)), as well as among speakers who reside in the Old

⁴⁶The measure of hierarchical orientation is based on the following questions in the World Values Survey (WVS): “Here is a list of qualities that children can be encouraged to learn at home. Which, if any, do you consider to be especially important?” Individuals are considered to have higher hierarchical orientation if they answered “Obedience” or “Religious Faith” and lower if they answered “Independence” or “Determination, Perseverance”. This measure is an inverse coding of the Autonomy Index introduced in (83).

⁴⁷The sample of languages in the WVS for which politeness distinctions data exists has little regional variation. This impedes the use of regional fixed-effects in the analysis of the main measure of politeness distinctions. Reassuringly, as established in Tables S9.9 and S9.14, exploiting the larger regional variations present in the additional measures of politeness distinctions, permits the analysis to account for regional fixed-effects, obtaining qualitatively similar results.

World and speak languages originated in the Old World (Columns (3) and (4)). These findings lend credence to the proposed hypothesis that the prevalence of hierarchical orientation and the presence of politeness distinctions have coevolved in the course of human history.

Additional Results

S7 The Origins of Periphrastic Future Tense

S7.A Crop Return and Periphrastic Future Tense

Table S7.1: Geographical Origins of Periphrastic Future Tense (OLS)

	Existence of Future Tense					
	(1)	(2)	(3)	(4)	(5)	(6)
Crop Return (pre-1500CE)	0.06** (0.03)	0.06** (0.03)	0.07** (0.03)	0.07** (0.03)	0.09*** (0.03)	0.11*** (0.03)
Absolute Latitude		-0.03 (0.03)	-0.03 (0.04)	0.03 (0.10)	0.05 (0.10)	0.15 (0.11)
Elevation		0.04 (0.04)	0.03 (0.04)	0.05 (0.05)	0.05 (0.04)	0.03 (0.05)
Ruggedness		-0.04 (0.04)	-0.03 (0.04)	-0.04 (0.04)	-0.05 (0.04)	-0.02 (0.04)
Coast Length		0.08*** (0.02)	0.07*** (0.02)	0.06*** (0.02)	0.06*** (0.02)	0.06*** (0.02)
Precipitation (mm/month)			0.01 (0.07)	0.02 (0.07)	0.01 (0.07)	0.01 (0.08)
Precipitation (mm/month) (std)			0.07** (0.03)	0.01 (0.05)	0.01 (0.05)	0.04 (0.05)
Precipitation Volatility			-0.04 (0.08)	-0.03 (0.08)	-0.03 (0.08)	-0.04 (0.08)
Precipitation Spatial Correlation			0.01 (0.04)	0.93*** (0.26)	0.91*** (0.26)	0.95*** (0.30)
Temperature (Daily Mean)				0.05 (0.07)	0.07 (0.06)	0.08 (0.08)
Temperature (Daily Mean) (std)				0.07 (0.05)	0.08* (0.05)	0.05 (0.04)
Temperature Volatility				-0.01 (0.08)	-0.06 (0.08)	-0.09 (0.09)
Temperature Spatial Correlation				-0.93*** (0.26)	-0.91*** (0.26)	-0.93*** (0.30)
Unproductive Period (pre-1500CE)					0.08** (0.03)	0.09*** (0.03)
Regional FE	No	No	No	No	No	Yes
Adjusted- R^2	0.01	0.02	0.02	0.06	0.07	0.11
Observations	275	275	275	275	275	275

Notes: This table documents the positive, statistically, and economically significant association between pre-1500CE potential crop return and the existence of periphrastic future tense in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All specifications in this table are identical to the ones presented in Table S4.1, but here the coefficients of the controls are presented. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S7.2: Crop Return and Periphrastic Future Tense

	Existence of Periphrastic Future Tense					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Probit						
Crop Return (pre-1500CE)	0.06** (0.03)	0.06** (0.03)	0.06** (0.03)	0.07** (0.03)	0.09*** (0.03)	0.12*** (0.03)
Geographical Controls	No	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes
Pseudo- R^2	0.01	0.03	0.04	0.08	0.10	0.14
Observations	275	275	275	275	275	275
Panel B: OLS - Spatial-Autocorrelation, Clustering and Selection On Unobservables						
Crop Return (pre-1500CE)	0.06** (0.03) ([0.04]) [0.04] {0.03}	0.06** (0.03) ([0.04]) [0.04] {0.03}	0.07** (0.03) ([0.03]) [0.03] {0.03}	0.07** (0.03) ([0.03]) [0.03] {0.03}	0.09*** (0.03) ([0.03]) [0.03] {0.03}	0.11*** (0.03) ([0.03]) [0.03] {0.03}
Altonji et al						-2.09
δ						-2.84
β -Oster						0.13
R^2	0.01	0.04	0.06	0.10	0.12	0.17

Notes: This table documents the positive, statistically, and economically significant association between pre-1500CE potential crop return and the existence of periphrastic future tense in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S4.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses, clustered at the language genus in parenthesis and squared brackets, spatial auto-correlation corrected standard errors (*101*) in squared brackets and Cliff-Ord ML in curly brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S7.3: Geographical Origins of the Future Tense
Robustness to Ancestry and Contact

	Existence of Future Tense								
	Genus-Level				Family-Level				Family + Genus
	Logit	RE Logit	PA Logit	MME Logit	Logit	RE Logit	PA Logit	MME Logit	MME Logit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Crop Return (pre-1500CE)	1.73*** (0.29)	1.97*** (0.48)	1.50*** (0.23)	2.17*** (0.51)	1.73*** (0.32)	1.34 (0.31)	1.54** (0.26)	1.97*** (0.50)	2.15** (0.66)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main Geographical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	275	275	275	275	275	275	275	275	275

Notes: Coefficients represent Odds ratios. RE Logit refers to random-effects logistic estimator, PA Logit refers to population-averaged logistic estimator, and MME Logit to multilevel mixed-effects logistic estimator. Heteroskedasticity robust standard error estimates clustered at the language family or genus level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S7.4: Long-Term Orientation and Tense

	Long-Term Orientation			
	Future	Past	Perfect	Any
	(1)	(2)	(3)	(4)
Crop Return (pre-1500CE)	0.29* (0.16)	0.45** (0.17)	0.46** (0.16)	0.38** (0.17)
Periphrastic Future Tense	0.15*** (0.04)			
Past tense		-0.07 (0.18)		
Perfect tense			-0.09 (0.10)	
Any inflectional tense				-0.21 (0.15)
Regional FE	Yes	Yes	Yes	Yes
Main Geographical Controls	Yes	Yes	Yes	Yes
Adjusted- R^2	0.24	0.17	0.18	0.27
Observations	106	58	58	88

Notes: Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests; All regressions include a constant.

Table S7.5: Agricultural Intensity and Crop Return

	Agricultural Intensity			
	Full Sample		Future Sample	
	(1)	(2)	(3)	(4)
Crop Return (pre-1500CE)	0.19*** (0.03)	0.22*** (0.02)	0.27*** (0.07)	0.30*** (0.06)
Regional FE	No	Yes	No	Yes
All Geographical Controls	No	Yes	No	Yes
Adjusted- R^2	0.04	0.64	0.07	0.61
Observations	1306	1306	264	264

Notes: This table documents the positive statistically and economically significant association between a language's contemporary homeland's crop return and the level of agricultural intensity of a pre-colonial society that speaks that language. Standardized coefficients. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

S8 Origins of Sex-Based Grammatical Gender

Table S8.1: Geographic Origins of Sex-Based Grammatical Gender (OLS)

	Existence of Sex-Based Grammatical Gender					
	(1)	(2)	(3)	(4)	(5)	(6)
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.13** (0.06)	-0.19*** (0.06)	-0.25*** (0.06)	-0.26*** (0.06)	-0.29*** (0.06)	-0.23*** (0.08)
All Crops (Average Caloric Yield, pre-1500)	0.17*** (0.06)	0.21*** (0.06)	0.28*** (0.06)	0.29*** (0.06)	0.32*** (0.06)	0.25*** (0.08)
Absolute Latitude		-0.10*** (0.04)	-0.19*** (0.04)	-0.06 (0.10)	-0.06 (0.10)	-0.13 (0.10)
Elevation		-0.01 (0.05)	-0.14*** (0.04)	-0.11** (0.05)	-0.11** (0.05)	-0.11** (0.05)
Ruggedness		-0.02 (0.05)	0.08 (0.05)	0.08 (0.05)	0.08 (0.05)	0.06 (0.05)
Coast Length		0.08*** (0.02)	0.09*** (0.02)	0.09*** (0.02)	0.08*** (0.02)	0.08*** (0.02)
Precipitation (mm/month)			-0.11* (0.07)	-0.11 (0.07)	-0.12 (0.07)	-0.08 (0.07)
Precipitation (mm/month) (std)			-0.02 (0.03)	0.03 (0.05)	0.04 (0.05)	0.03 (0.05)
Precipitation Volatility			0.03 (0.07)	-0.01 (0.08)	0.00 (0.08)	0.01 (0.08)
Precipitation Spatial Correlation			0.13*** (0.03)	0.14 (0.28)	0.15 (0.29)	0.12 (0.28)
Temperature (Daily Mean)				0.08 (0.07)	0.07 (0.07)	0.04 (0.07)
Temperature (Daily Mean) (std)				-0.07 (0.05)	-0.08* (0.04)	-0.07 (0.05)
Temperature Volatility				-0.08 (0.09)	-0.07 (0.09)	-0.10 (0.09)
Temperature Spatial Correlation				-0.01 (0.29)	-0.02 (0.29)	-0.02 (0.28)
Unproductive Period (pre-1500CE)					-0.06* (0.03)	-0.02 (0.03)
Regional FE	No	No	No	No	No	Yes
Adjusted- R^2	0.03	0.06	0.14	0.15	0.15	0.21
Observations	217	217	217	217	217	217

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of sex-based grammatical gender in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S5.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.2: Geographic Origins of Sex-Based Grammatical Gender

	Existence of Sex-Based Grammatical Gender					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Probit						
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.13** (0.05)	-0.19*** (0.06)	-0.25*** (0.06)	-0.26*** (0.06)	-0.28*** (0.06)	-0.20** (0.08)
All Crops (Average Caloric Yield, pre-1500)	0.16*** (0.05)	0.21*** (0.06)	0.28*** (0.06)	0.30*** (0.06)	0.32*** (0.06)	0.22*** (0.07)
Geographical Controls	No	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes
Pseudo- R^2	0.03	0.07	0.16	0.18	0.18	0.25
Observations	216	216	216	216	216	216
Panel B: OLS - Spatial-Autocorrelation, Clustering and Selection On Unobservables						
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.13** (0.06) ([0.07]) [0.06] {0.06}	-0.19*** (0.06) ([0.07]) [0.07] {0.06}	-0.25*** (0.06) ([0.07]) [0.07] {0.06}	-0.26*** (0.06) ([0.07]) [0.07] {0.06}	-0.29*** (0.06) ([0.07]) [0.07] {0.07}	-0.23*** (0.08) ([0.08]) [0.07] {0.09}
Altonji et al						-2.26
δ						-1.29
β -Oster						-0.27
All Crops (Average Caloric Yield, pre-1500)	0.17*** (0.06) ([0.07]) [0.06] {0.06}	0.21*** (0.06) ([0.07]) [0.06] {0.06}	0.28*** (0.06) ([0.07]) [0.06] {0.06}	0.29*** (0.06) ([0.07]) [0.06] {0.06}	0.32*** (0.06) ([0.07]) [0.06] {0.06}	0.25*** (0.08) ([0.07]) [0.07] {0.08}
Altonji et al						-3.02
δ						-1.56
β -Oster						0.28
R^2	0.04	0.09	0.18	0.20	0.21	0.28

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of sex-based grammatical gender in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S7.2. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses, clustered at the language genus in parenthesis and squared brackets, spatial auto-correlation corrected standard errors (101) in squared brackets and Cliff-Ord ML in curly brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.3: Geographic Origins of Sex-Based Grammatical Gender
Robustness to Ancestry and Contact

	Existence of Sex-Based Grammatical Gender								
	Genus-Level				Family-Level				Family + Genus
	Logit	RE Logit	PA Logit	MME Logit	Logit	RE Logit	PA Logit	MME Logit	MME Logit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Plow Negative Crops	0.28**	0.00*	0.49	0.07**	0.28***	0.42	0.56	0.15	0.09
(Average Caloric Yield, pre-1500)	(0.15)	(0.01)	(0.21)	(0.09)	(0.14)	(0.71)	(0.21)	(0.22)	(0.23)
All Crops	3.93***	501.35**	2.04**	12.05**	3.93***	7.87	2.04**	16.21**	44.94
(Average Caloric Yield, pre-1500)	(1.73)	(1365.42)	(0.70)	(12.02)	(1.51)	(10.61)	(0.63)	(21.01)	(110.15)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main Geographical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	217	217	217	217	217	217	217	217	217

Notes: Coefficients represent Odds ratios. RE Logit refers to random-effects logistic estimator, PA Logit refers to population-averaged logistic estimator, and MME Logit to multilevel mixed-effects logistic estimator. Heteroskedasticity robust standard error estimates clustered at the language family or genus level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

S8.A Alternative Measure for Sex-Based Grammatical Gender (Extensive Measure)

Table S8.4: Geographic Origins of Sex-Based Grammatical Gender (OLS)

	Existence of Sex-Based Grammatical Gender					
	(1)	(2)	(3)	(4)	(5)	(6)
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.03 (0.06)	-0.10 (0.06)	-0.16** (0.06)	-0.17*** (0.06)	-0.23*** (0.07)	-0.22*** (0.08)
All Crops (Average Caloric Yield, pre-1500)	0.09 (0.05)	0.14** (0.06)	0.21*** (0.06)	0.22*** (0.06)	0.27*** (0.06)	0.26*** (0.07)
Absolute Latitude		-0.10*** (0.04)	-0.20*** (0.04)	-0.09 (0.10)	-0.10 (0.10)	-0.15 (0.09)
Elevation		0.02 (0.05)	-0.11** (0.04)	-0.09* (0.05)	-0.11** (0.05)	-0.13*** (0.04)
Ruggedness		-0.05 (0.05)	0.04 (0.05)	0.04 (0.05)	0.05 (0.05)	0.05 (0.05)
Coast Length		0.07*** (0.02)	0.08*** (0.02)	0.08*** (0.02)	0.07*** (0.02)	0.07*** (0.02)
Precipitation (mm/month)			-0.09 (0.06)	-0.10 (0.07)	-0.11 (0.07)	-0.06 (0.07)
Precipitation (mm/month) (std)			-0.03 (0.03)	0.01 (0.05)	0.01 (0.05)	0.00 (0.05)
Precipitation Volatility			-0.00 (0.07)	-0.02 (0.08)	-0.00 (0.08)	0.02 (0.07)
Precipitation Spatial Correlation			0.13*** (0.03)	0.20 (0.28)	0.22 (0.28)	0.18 (0.26)
Temperature (Daily Mean)				0.03 (0.07)	0.02 (0.07)	-0.02 (0.07)
Temperature (Daily Mean) (std)				-0.06 (0.05)	-0.07 (0.04)	-0.06 (0.04)
Temperature Volatility				-0.08 (0.09)	-0.06 (0.09)	-0.08 (0.09)
Temperature Spatial Correlation				-0.07 (0.29)	-0.06 (0.29)	-0.06 (0.26)
Unproductive Period (pre-1500CE)					-0.08*** (0.03)	-0.03 (0.03)
Regional FE	No	No	No	No	No	Yes
Adjusted- R^2	0.01	0.04	0.13	0.13	0.15	0.24
Observations	245	245	245	245	245	245

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of sex-based grammatical gender in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S5.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.5: Geographic Origins of Sex-Based Grammatical Gender

	Existence of Sex-Based Grammatical Gender					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Probit						
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.03 (0.05)	-0.10* (0.06)	-0.16** (0.06)	-0.17*** (0.06)	-0.22*** (0.06)	-0.21*** (0.07)
All Crops (Average Caloric Yield, pre-1500)	0.09 (0.05)	0.14** (0.06)	0.21*** (0.06)	0.22*** (0.06)	0.26*** (0.06)	0.24*** (0.07)
Geographical Controls	No	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes
Pseudo- R^2	0.01	0.05	0.14	0.15	0.17	0.26
Observations	245	245	245	245	245	245
Panel B: OLS - Spatial-Autocorrelation, Clustering and Selection On Unobservables						
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.03 (0.06) ([0.07]) [0.06] {0.05}	-0.10 (0.06) ([0.08]) [0.07] {0.06}	-0.16** (0.06) ([0.08]) [0.07] {0.06}	-0.17*** (0.06) ([0.07]) [0.07] {0.06}	-0.23*** (0.07) ([0.07]) [0.07] {0.06}	-0.22*** (0.08) ([0.07]) [0.07] {0.08}
Altonji et al						-1.14
δ						-0.89
β -Oster						-0.29
All Crops (Average Caloric Yield, pre-1500)	0.09 (0.05) ([0.07]) [0.06] {0.05}	0.14** (0.06) ([0.07]) [0.06] {0.06}	0.21*** (0.06) ([0.07]) [0.06] {0.06}	0.22*** (0.06) ([0.07]) [0.06] {0.06}	0.27*** (0.06) ([0.07]) [0.06] {0.06}	0.26*** (0.07) ([0.07]) [0.07] {0.07}
Altonji et al						-1.51
δ						-0.89
β -Oster						0.31
R^2	0.02	0.07	0.17	0.18	0.20	0.31

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of sex-based grammatical gender in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S7.2. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses, clustered at the language genus in parenthesis and squared brackets, spatial auto-correlation corrected standard errors (101) in squared brackets and Cliff-Ord ML in curly brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.6: Geographic Origins of Sex-Based Grammatical Gender
Robustness to Ancestry and Contact

	Existence of Sex-Based Grammatical Gender								
	Genus-Level				Family-Level				Family + Genus
	Logit	RE Logit	PA Logit	MME Logit	Logit	RE Logit	PA Logit	MME Logit	MME Logit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Plow Negative Crops	0.30***	0.01**	0.57	0.12**	0.30***	0.20	0.51*	0.13**	0.06
(Average Caloric Yield, pre-1500)	(0.14)	(0.01)	(0.21)	(0.13)	(0.12)	(0.23)	(0.18)	(0.13)	(0.18)
All Crops	3.92***	251.44***	1.76*	9.53**	3.92***	12.55***	2.38***	19.70***	69.95
(Average Caloric Yield, pre-1500)	(1.63)	(526.64)	(0.51)	(8.37)	(1.48)	(12.23)	(0.73)	(20.62)	(198.79)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main Geographical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	245	245	245	245	245	245	245	245	245

Notes: Coefficients represent Odds ratios. RE Logit refers to random-effects logistic estimator, PA Logit refers to population-averaged logistic estimator, and MME Logit to multilevel mixed-effects logistic estimator. Heteroskedasticity robust standard error estimates clustered at the language family or genus level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.7: Persistent Impact of Geographical Characteristics on Gender:
Linguistic Homeland vs. Urheimat
(Languages Outside Urheimat)

	Existence of Sex-Based Gender System					
	Migratory Distance to Urheimat					
	Any Distance			At Least 1 Week		
	Homeland	Urheimat	Both	Homeland	Urheimat	Both
	(1)	(2)	(3)	(4)	(5)	(6)
Homeland's Plow Negative Crops (Avg. Caloric Yield, pre-1500)	-0.22*** (0.06)		-0.08 (0.07)	-0.11 (0.08)		-0.10* (0.06)
Homeland's Crops (Average Caloric Yield, pre-1500)	0.25*** (0.06)		0.14** (0.06)	0.13 (0.09)		0.11* (0.06)
Urheimat's Plow Negative Crops (Avg. Caloric Yield, pre-1500)		-0.31*** (0.09)	-0.27*** (0.10)		-0.52*** (0.17)	-0.54** (0.20)
Urheimat's Crops (Average Caloric Yield, pre-1500)		0.50*** (0.10)	0.46*** (0.10)		0.71*** (0.21)	0.71*** (0.24)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes
Homeland Geographical Characteristics	Yes	No	Yes	Yes	No	Yes
Urheimat Geographical Characteristics	No	Yes	Yes	No	Yes	Yes
Adjusted- R^2	0.26	0.46	0.45	0.34	0.57	0.54
Observations	222	222	222	155	155	155
Language Families	57	57	57	30	30	30

Notes: This table explores the relative contributions of agricultural productivity in the contemporary homeland vs. the Urheimat to the presence of sex-based grammatical gender in a daughter language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

S8.B Alternative Measure for Sex-Based Grammatical Gender (Intensive Measure)

Table S8.8: Geographic Origins of Intensity of Sex-Based Grammatical Gender System (OLS)

	Intensity of Sex-Based Grammatical Gender System					
	(1)	(2)	(3)	(4)	(5)	(6)
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.25*** (0.09)	-0.32*** (0.10)	-0.40*** (0.11)	-0.39*** (0.11)	-0.47*** (0.12)	-0.36** (0.14)
All Crops (Average Caloric Yield, pre-1500)	0.26** (0.10)	0.32*** (0.11)	0.41*** (0.12)	0.39*** (0.12)	0.47*** (0.14)	0.40*** (0.14)
Absolute Latitude		-0.04 (0.04)	-0.21*** (0.07)	0.15 (0.16)	0.14 (0.16)	0.04 (0.15)
Elevation		0.06 (0.08)	-0.10 (0.09)	-0.02 (0.08)	-0.03 (0.08)	-0.07 (0.08)
Ruggedness		-0.14* (0.07)	-0.01 (0.07)	-0.04 (0.09)	-0.03 (0.08)	-0.00 (0.08)
Coast Length		0.03 (0.05)	0.03 (0.05)	0.03 (0.06)	0.02 (0.06)	0.04 (0.06)
Precipitation (mm/month)			-0.29** (0.11)	-0.25** (0.11)	-0.28** (0.11)	-0.20* (0.10)
Precipitation (mm/month) (std)			0.03 (0.04)	-0.01 (0.08)	-0.00 (0.08)	0.01 (0.08)
Precipitation Volatility			0.07 (0.11)	0.03 (0.13)	0.06 (0.12)	0.08 (0.12)
Precipitation Spatial Correlation			0.09** (0.05)	-0.09 (0.43)	-0.02 (0.45)	0.08 (0.47)
Temperature (Daily Mean)				0.22* (0.12)	0.20 (0.14)	0.08 (0.13)
Temperature (Daily Mean) (std)				0.03 (0.08)	0.02 (0.08)	-0.02 (0.08)
Temperature Volatility				-0.20 (0.14)	-0.17 (0.14)	-0.13 (0.14)
Temperature Spatial Correlation				0.17 (0.43)	0.13 (0.44)	0.00 (0.46)
Unproductive Period (pre-1500CE)					-0.14** (0.07)	-0.07 (0.06)
Regional FE	No	No	No	No	No	Yes
Adjusted- R^2	0.03	0.03	0.06	0.07	0.09	0.17
Observations	181	181	181	181	181	181

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of sex-based grammatical gender in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S5.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.9: Geographic Origins of Intensity of Sex-Based Grammatical Gender System

	Intensity of Sex-Based Grammatical Gender System					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS - Spatial-Autocorrelation, Clustering and Selection On Unobservables						
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.25*** (0.09) ([0.11]) [0.10] {0.10}	-0.32*** (0.10) ([0.12]) [0.10] {0.11}	-0.40*** (0.11) ([0.14]) [0.11] {0.12}	-0.39*** (0.11) ([0.13]) [0.11] {0.12}	-0.47*** (0.12) ([0.15]) [0.12] {0.12}	-0.36** (0.14) ([0.14]) [0.12] {0.15}
Altonji et al						-3.22
δ						-1.95
β -Oster						-0.40
All Crops (Average Caloric Yield, pre-1500)	0.26** (0.10) ([0.12]) [0.11] {0.10}	0.32*** (0.11) ([0.13]) [0.11] {0.11}	0.41*** (0.12) ([0.14]) [0.12] {0.11}	0.39*** (0.12) ([0.14]) [0.11] {0.11}	0.47*** (0.14) ([0.15]) [0.12] {0.12}	0.40*** (0.14) ([0.14]) [0.12] {0.14}
Altonji et al						-2.85
δ						-1.36
β -Oster						0.45
R^2	0.04	0.06	0.12	0.14	0.16	0.26
Observations	181	181	181	181	181	181

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of sex-based grammatical gender in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S7.2. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses, clustered at the language genus in parenthesis and squared brackets, spatial auto-correlation corrected standard errors (*101*) in squared brackets and Cliff-Ord ML in curly brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.10: Persistent Impact of Geographical Characteristics on Intensity of Sex-Based Grammatical Gender System
Linguistic Homeland vs. Urheimat
(Languages Outside Urheimat)

	Existence of Sex-Based Gender System					
	Migratory Distance to Urheimat					
	Any Distance			At Least 1 Week		
	Homeland	Urheimat	Both	Homeland	Urheimat	Both
	(1)	(2)	(3)	(4)	(5)	(6)
Homeland's Plow Negative Crops (Avg. Caloric Yield, pre-1500)	-0.39** (0.18)		-0.08 (0.13)	-0.31 (0.23)		-0.39 (0.26)
Homeland's Crops (Average Caloric Yield, pre-1500)	0.42** (0.20)		0.17 (0.14)	0.30 (0.23)		0.22 (0.19)
Urheimat's Plow Negative Crops (Avg. Caloric Yield, pre-1500)		-0.43*** (0.11)	-0.47*** (0.13)		-0.61 (0.41)	-0.42 (0.39)
Urheimat's Crops (Average Caloric Yield, pre-1500)		0.55*** (0.12)	0.58*** (0.14)		0.73* (0.39)	0.60 (0.42)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes
Homeland Geographical Characteristics	Yes	No	Yes	Yes	No	Yes
Urheimat Geographical Characteristics	No	Yes	Yes	No	Yes	Yes
Adjusted- R^2	0.17	0.34	0.31	0.11	0.42	0.42
Observations	162	162	162	109	109	109
Language Families	55	55	55	30	30	30

Notes: This table explores the relative contributions of agricultural productivity in the contemporary homeland vs. the Urheimat to the presence of sex-based grammatical gender in a daughter language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.11: Geographic Origins of Usage of the Plow and Intensive Sex-Based Grammatical Genders

	Mechanism				
	Plow		Grammatical Gender		
	OLS	OLS	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.26*** (0.04)	-0.06* (0.03)			
Average Caloric Yield (All Crops, pre-1500)	0.26*** (0.04)	0.10*** (0.03)			
Aboriginal Plow			0.65** (0.31)	0.56 (0.37)	0.72 (0.89)
All Geographic Controls	No	Yes	No	Yes	Yes
Regional FE	No	Yes	No	Yes	Yes
First-stage F-statistic					10.33
Adjusted- R^2	0.19	0.47	0.09	0.21	0.21
Observations	1158	1158	133	133	133

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of sex-based grammatical gender in a language. The first two columns provide the association between the geographical determinants of plow suitability and actual usage of the plow, and columns (3) and (4) provide evidence on the association between actual usage of the plow and the emergence of sex-based grammatical gender. Column (5) tests the mechanism through which the usage of the plow affects the existence of sex-based grammatical gender in a language using an IV strategy. Heteroskedasticity robust standard error estimates clustered at the genus level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

S8.C Origins of Gender Distinctions in Pronouns

Table S8.12: Geographic Origins of Gender Distinctions in Pronouns (OLS)

	Existence of Gender Distinctions in Pronouns					
	(1)	(2)	(3)	(4)	(5)	(6)
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.08*	-0.12**	-0.13***	-0.15***	-0.17***	-0.18***
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
All Crops (Average Caloric Yield, pre-1500)	0.10**	0.14***	0.15***	0.18***	0.19***	0.21***
	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)
Absolute Latitude		-0.04	-0.08**	0.06	0.05	0.01
		(0.03)	(0.03)	(0.07)	(0.07)	(0.08)
Elevation		0.03	-0.03	-0.01	-0.02	-0.03
		(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Ruggedness		-0.11***	-0.05	-0.07*	-0.07*	-0.06*
		(0.03)	(0.03)	(0.04)	(0.04)	(0.04)
Coast Length		0.07***	0.07***	0.07***	0.07***	0.08***
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Precipitation (mm/month)			0.02	-0.01	-0.02	0.01
			(0.05)	(0.06)	(0.06)	(0.06)
Precipitation (mm/month) (std)			-0.00	0.02	0.02	0.01
			(0.03)	(0.04)	(0.04)	(0.04)
Precipitation Volatility			-0.09*	-0.09*	-0.09	-0.08
			(0.05)	(0.05)	(0.05)	(0.05)
Precipitation Spatial Correlation			0.04	0.24	0.26	0.34
			(0.03)	(0.23)	(0.23)	(0.25)
Temperature (Daily Mean)				0.01	0.01	-0.03
				(0.06)	(0.06)	(0.06)
Temperature (Daily Mean) (std)				-0.04	-0.04	-0.04
				(0.04)	(0.04)	(0.04)
Temperature Volatility				-0.16***	-0.16**	-0.19***
				(0.06)	(0.06)	(0.06)
Temperature Spatial Correlation				-0.19	-0.20	-0.29
				(0.23)	(0.23)	(0.25)
Unproductive Period (pre-1500CE)					-0.03	-0.02
					(0.03)	(0.03)
Regional FE	No	No	No	No	No	Yes
Adjusted- R^2	0.01	0.06	0.08	0.10	0.10	0.13
Observations	354	354	354	354	354	354

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of gender distinctions in independent personal pronouns in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S5.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.13: Geographic Origins of Gender Distinctions in Pronouns (Probit)

	Existence of Gender Distinctions in Pronouns					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Probit						
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.06 (0.04)	-0.10** (0.05)	-0.11** (0.05)	-0.13*** (0.05)	-0.14*** (0.05)	-0.16** (0.06)
All Crops (Average Caloric Yield, pre-1500)	0.09** (0.04)	0.12*** (0.05)	0.14*** (0.05)	0.16*** (0.05)	0.17*** (0.05)	0.19*** (0.06)
Geographical Controls	No	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes
Pseudo- R^2	0.01	0.07	0.09	0.11	0.11	0.15
Observations	350	350	350	350	350	350
Panel B: OLS - Spatial-Autocorrelation, Clustering and Selection On Unobservables						
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.08* (0.04) ([0.05]) [0.05] {0.04}	-0.12** (0.05) ([0.06]) [0.05] {0.05}	-0.13*** (0.05) ([0.05]) [0.05] {0.05}	-0.15*** (0.05) ([0.05]) [0.05] {0.05}	-0.17*** (0.05) ([0.05]) [0.05] {0.05}	-0.18*** (0.06) ([0.07]) [0.06] {0.07}
Altonji et al						-1.77
δ						-1.12
β -Oster						-0.21
All Crops (Average Caloric Yield, pre-1500)	0.10** (0.04) ([0.05]) [0.05] {0.04}	0.14*** (0.05) ([0.06]) [0.05] {0.05}	0.15*** (0.05) ([0.05]) [0.05] {0.05}	0.18*** (0.05) ([0.05]) [0.04] {0.05}	0.19*** (0.05) ([0.05]) [0.05] {0.05}	0.21*** (0.06) ([0.06]) [0.06] {0.06}
Altonji et al						-1.89
δ						-0.90
β -Oster						0.24
R^2	0.01	0.08	0.11	0.13	0.14	0.18

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of gender distinctions in pronouns in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S7.2. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses, clustered at the language genus in parenthesis and squared brackets, spatial auto-correlation corrected standard errors (*101*) in squared brackets and Cliff-Ord ML in curly brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.14: Geographic Origins of Gender Distinctions in Pronouns
Robustness to Ancestry and Contact

	Existence of Gender Distinctions in Pronouns								
	Genus-Level				Family-Level				Family + Genus
	Logit	RE Logit	PA Logit	MME Logit	Logit	RE Logit	PA Logit	MME Logit	MME Logit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Plow Negative Crops	0.40**	0.24**	0.44**	0.23**	0.40**	0.37**	0.55**	0.35*	0.25*
(Average Caloric Yield, pre-1500)	(0.15)	(0.16)	(0.15)	(0.14)	(0.17)	(0.18)	(0.16)	(0.20)	(0.19)
All Crops	2.96***	5.85***	2.73***	5.43***	2.96***	3.78***	2.13***	3.89***	5.66**
(Average Caloric Yield, pre-1500)	(1.01)	(3.68)	(0.85)	(3.04)	(1.12)	(1.72)	(0.56)	(1.98)	(3.95)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main Geographical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	351	351	351	351	351	351	351	351	351

Notes: Coefficients represent Odds ratios. RE Logit refers to random-effects logistic estimator, PA Logit refers to population-averaged logistic estimator, and MME Logit to multilevel mixed-effects logistic estimator. Heteroskedasticity robust standard error estimates clustered at the language family or genus level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.15: Persistent Impact of Geographical Characteristics on Gender Distinctions in Pronouns
Linguistic Homeland vs. Urheimat
(Languages Outside Urheimat)

	Existence of Gender Distinctions in Independent Personal Pronouns					
	Migratory Distance to Urheimat					
	Any Distance			At Least 1 Week		
	Homeland	Urheimat	Both	Homeland	Urheimat	Both
	(1)	(2)	(3)	(4)	(5)	(6)
Homeland's Plow Negative Crops (Avg. Caloric Yield, pre-1500)	-0.18** (0.07)		-0.05 (0.07)	-0.11 (0.10)		-0.07 (0.08)
Homeland's Crops (Average Caloric Yield, pre-1500)	0.20*** (0.07)		0.11** (0.05)	0.12 (0.09)		0.08 (0.07)
Urheimat's Plow Negative Crops (Avg. Caloric Yield, pre-1500)		-0.22*** (0.07)	-0.22*** (0.06)		-0.21** (0.08)	-0.20** (0.08)
Urheimat's Crops (Average Caloric Yield, pre-1500)		0.22*** (0.07)	0.22*** (0.06)		0.23*** (0.08)	0.22*** (0.07)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes
Homeland Geographical Characteristics	Yes	No	Yes	Yes	No	Yes
Urheimat Geographical Characteristics	No	Yes	Yes	No	Yes	Yes
Adjusted- R^2	0.15	0.25	0.27	0.17	0.34	0.37
Observations	316	316	316	207	207	207
Language Families	76	76	76	36	36	36

Notes: This table explores the relative contributions of agricultural productivity in the contemporary homeland vs. the Urheimat to the presence of sex-based grammatical gender in a daughter language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S8.16: Geographic Origins of Usage of the Plow and Gender Distinctions in Pronouns

	Mechanism				
	Plow		Gender in Pronouns		
	OLS	OLS	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)
Plow Negative Crops (Average Caloric Yield, pre-1500)	-0.26*** (0.04)	-0.06* (0.03)			
Average Caloric Yield (All Crops, pre-1500)	0.26*** (0.04)	0.10*** (0.03)			
Aboriginal Plow			0.24** (0.10)	0.22* (0.12)	1.18*** (0.44)
All Geographic Controls	No	Yes	No	Yes	Yes
Regional FE	No	Yes	No	Yes	Yes
First-stage F-statistic					10.89
Adjusted- R^2	0.19	0.47	0.04	0.10	-0.23
Observations	1158	1158	220	220	220

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of gender distinctions in independent personal pronouns in a language. The first two columns provide the association between the geographical determinants of plow suitability and actual usage of the plow, and columns (3) and (4) provide evidence on the association between actual usage of the plow and the emergence of sex-based grammatical gender. Column (5) shows the mechanism of the plow on the existence of gender in pronouns using an IV strategy. Heteroskedasticity robust standard error estimates clustered at the genus level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

S9 Origins of Politeness Distinctions

Table S9.1: Geographic Origins of Politeness Distinctions (OLS)

	Existence of Politeness Distinctions					
	(1)	(2)	(3)	(4)	(5)	(6)
Ecological Diversity	0.15*** (0.03)	0.10*** (0.04)	0.10*** (0.04)	0.10** (0.04)	0.10** (0.04)	0.10** (0.04)
All Crops (Average Caloric Yield, pre-1500)		0.11*** (0.03)	0.10*** (0.03)	0.11*** (0.03)	0.11*** (0.03)	0.12*** (0.03)
Absolute Latitude		0.07** (0.03)	0.09** (0.04)	0.23** (0.11)	0.24** (0.11)	0.07 (0.11)
Elevation		0.01 (0.05)	0.05 (0.05)	0.09* (0.05)	0.09* (0.05)	0.10** (0.05)
Ruggedness		-0.04 (0.04)	-0.06* (0.04)	-0.07* (0.04)	-0.06* (0.04)	-0.09** (0.04)
Coast Length		0.07* (0.04)	0.06 (0.04)	0.06 (0.05)	0.06 (0.05)	0.04 (0.05)
Precipitation (mm/month)			0.03 (0.06)	0.03 (0.06)	0.04 (0.06)	-0.02 (0.06)
Precipitation (mm/month) (std)			0.04 (0.04)	0.05 (0.07)	0.06 (0.07)	0.07 (0.07)
Precipitation Volatility			-0.06 (0.07)	-0.07 (0.07)	-0.08 (0.06)	-0.04 (0.07)
Precipitation Spatial Correlation			-0.08** (0.04)	0.31 (0.28)	0.31 (0.29)	0.35 (0.27)
Temperature (Daily Mean)				0.07 (0.08)	0.07 (0.08)	0.04 (0.07)
Temperature (Daily Mean) (std)				-0.01 (0.07)	-0.03 (0.07)	-0.07 (0.06)
Temperature Volatility				-0.09 (0.09)	-0.07 (0.08)	0.03 (0.08)
Temperature Spatial Correlation				-0.39 (0.29)	-0.39 (0.30)	-0.42 (0.28)
Unproductive Period (pre-1500CE)					-0.06* (0.04)	-0.07* (0.04)
Regional FE	No	No	No	No	No	Yes
Adjusted- R^2	0.09	0.18	0.20	0.20	0.21	0.32
Observations	198	198	198	198	198	198

Notes: This table documents the positive, statistically, and economically significant association between ecological diversity and the existence of politeness distinctions in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S7.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S9.2: Geographic Origins of Politeness Distinctions

	Existence of Politeness Distinctions					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Probit						
Ecological Diversity	0.14*** (0.03)	0.09*** (0.03)	0.09*** (0.03)	0.10*** (0.03)	0.10*** (0.03)	0.10*** (0.03)
Geographical Controls	No	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes
Pseudo- R^2	0.07	0.21	0.23	0.24	0.25	0.36
Observations	180	180	180	180	180	180
Panel B: OLS - Spatial-Autocorrelation, Clustering and Selection On Unobservables						
Ecological Diversity	0.15*** (0.03) ([0.04]) [0.03] {0.03}	0.10*** (0.04) ([0.04]) [0.04] {0.03}	0.10*** (0.04) ([0.04]) [0.04] {0.03}	0.10** (0.04) ([0.04]) [0.04] {0.03}	0.10** (0.04) ([0.04]) [0.04] {0.03}	0.10** (0.04) ([0.03]) [0.04] {0.03}
Altonji et al						3.10
δ						5.02
β -Oster						0.08
R^2	0.10	0.21	0.24	0.26	0.27	0.39

Notes: This table documents the positive, statistically, and economically significant association between ecological diversity and the existence of politeness distinctions in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S4.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses, clustered at the language genus in parenthesis and squared brackets, spatial auto-correlation corrected standard errors (*101*) in squared brackets and Cliff-Ord ML in curly brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S9.3: Geographical Origins of Politeness Distinctions
Robustness to Ancestry and Contact

	Existence of Politeness Distinctions								
	Genus-Level				Family-Level				Family + Genus
	Logit	RE Logit	PA Logit	MME Logit	Logit	RE Logit	PA Logit	MME Logit	MME Logit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ecological Diversity	2.04*** (0.55)	2.04*** (0.51)	1.99*** (0.49)	2.55** (0.98)	2.04*** (0.46)	2.06*** (0.54)	2.07*** (0.52)	2.10*** (0.53)	2.62*** (0.88)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main Geographical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	180	180	180	180	180	180	180	180	180

Notes: Coefficients represent Odds ratios. RE Logit refers to random-effects logistic estimator, PA Logit refers to population-averaged logistic estimator, and MME Logit to multilevel mixed-effects logistic estimator. Heteroskedasticity robust standard error estimates clustered at the language family or genus level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

S9.A Alternative Measure of Politeness Distinctions (Extensive Margin)

Table S9.4: Geographic Origins of Politeness Distinctions (Alternative measure - OLS)

	Existence of politeness Distinctions (Alt. Measure)					
	(1)	(2)	(3)	(4)	(5)	(6)
Ecological Diversity	0.08*** (0.03)	0.09*** (0.03)	0.07** (0.03)	0.09*** (0.03)	0.09*** (0.03)	0.08*** (0.03)
All Crops (Average Caloric Yield, pre-1500)		0.03 (0.02)	0.04 (0.02)	0.04 (0.02)	0.03 (0.02)	0.06** (0.03)
Absolute Latitude		-0.03 (0.02)	-0.04 (0.03)	-0.05 (0.06)	-0.05 (0.06)	-0.10 (0.07)
Elevation		-0.06** (0.02)	-0.08*** (0.03)	-0.07* (0.04)	-0.07* (0.04)	-0.06* (0.03)
Ruggedness		0.01 (0.02)	0.02 (0.02)	0.03 (0.02)	0.04 (0.02)	-0.00 (0.03)
Coast Length		0.01 (0.03)	0.01 (0.03)	0.01 (0.03)	0.02 (0.03)	0.02 (0.03)
Precipitation (mm/month)			-0.04 (0.03)	-0.05 (0.05)	-0.05 (0.04)	-0.07 (0.04)
Precipitation (mm/month) (std)			0.06* (0.03)	0.12** (0.05)	0.12** (0.06)	0.12** (0.05)
Precipitation Volatility			0.01 (0.04)	-0.00 (0.04)	-0.01 (0.04)	-0.01 (0.04)
Precipitation Spatial Correlation			0.02 (0.02)	0.05 (0.19)	0.05 (0.19)	0.00 (0.19)
Temperature (Daily Mean)				-0.03 (0.07)	-0.03 (0.06)	-0.06 (0.06)
Temperature (Daily Mean) (std)				-0.08 (0.05)	-0.09* (0.05)	-0.11** (0.05)
Temperature Volatility				-0.01 (0.05)	0.00 (0.05)	0.06 (0.05)
Temperature Spatial Correlation				-0.03 (0.19)	-0.04 (0.19)	0.01 (0.19)
Unproductive Period (pre-1500CE)					-0.03* (0.02)	-0.04* (0.02)
Regional FE	No	No	No	No	No	Yes
Adjusted- R^2	0.06	0.09	0.09	0.09	0.10	0.25
Observations	198	198	198	198	198	198

Notes: This table documents the positive, statistically, and economically significant association between ecological diversity and the existence of politeness distinctions in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S7.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S9.5: Geographic Origins of Politeness Distinctions (Alternative measure - Probit)

	Existence of Politeness Distinctions (Alternative Measure)					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Probit						
Ecological Diversity	0.09*** (0.02)	0.11*** (0.03)	0.08*** (0.02)	0.08*** (0.02)	0.08*** (0.02)	0.07*** (0.02)
Geographical Controls	No	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes
Pseudo- R^2	0.09	0.21	0.26	0.27	0.29	0.51
Observations	140	140	140	140	140	140
Panel B: OLS - Spatial-Autocorrelation, Clustering and Selection On Unobservables						
Ecological Diversity	0.08*** (0.03) ([0.04]) [0.03] {0.02}	0.09*** (0.03) ([0.05]) [0.04] {0.02}	0.07** (0.03) ([0.04]) [0.04] {0.02}	0.09*** (0.03) ([0.05]) [0.04] {0.02}	0.09*** (0.03) ([0.05]) [0.04] {0.02}	0.08*** (0.03) ([0.03]) [0.03] {0.02}
Altonji et al						-7.30
δ						-12.31
β -Oster						0.09
R^2	0.07	0.11	0.14	0.16	0.17	0.33
Observations	198	198	198	198	198	198

Notes: This table documents the positive, statistically, and economically significant association between ecological diversity and the existence of politeness distinctions in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S4.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses, clustered at the language genus in parenthesis and squared brackets, spatial auto-correlation corrected standard errors (*101*) in squared brackets and Cliff-Ord ML in curly brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S9.6: Geographical Origins of Politeness Distinctions (Alternative measure)
Robustness to Ancestry and Contact

	Existence of Politeness Distinctions								
	Genus-Level				Family-Level				Family + Genus
	Logit	RE Logit	PA Logit	MME Logit	Logit	RE Logit	PA Logit	MME Logit	MME Logit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ecological Diversity	3.54** (2.27)	3.54*** (1.71)	3.51*** (1.61)	4.75 (5.22)	3.54* (2.34)	3.54*** (1.71)	3.70*** (1.49)	3.63** (2.28)	4.21 (4.05)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Main Geographical Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	140	140	140	140	140	140	140	140	140

Notes: Coefficients represent Odds ratios. RE Logit refers to random-effects logistic estimator, PA Logit refers to population-averaged logistic estimator, and MME Logit to multilevel mixed-effects logistic estimator. Heteroskedasticity robust standard error estimates clustered at the language family or genus level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S9.7: Persistent Impact of Geographical Characteristics on Politeness (Alternative Measure)
Linguistic Homeland vs. Urheimat
(Languages Outside Urheimat)

	Existence of Politeness Distinctions					
	Migratory Distance to Urheimat					
	Any Distance			At Least 1 Week		
	Homeland	Urheimat	Both	Homeland	Urheimat	Both
	(1)	(2)	(3)	(4)	(5)	(6)
Homeland's Ecological Diversity	0.10** (0.04)		0.10** (0.04)	0.10** (0.04)		0.08** (0.03)
Urheimat's Ecological Diversity		-0.02 (0.03)	-0.06* (0.03)		0.05 (0.05)	-0.01 (0.04)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes
Homeland Geographical Characteristics	Yes	No	Yes	Yes	No	Yes
Urheimat Geographical Characteristics	No	Yes	Yes	No	Yes	Yes
Adjusted- R^2	0.27	0.14	0.30	0.34	0.23	0.41
Observations	180	180	180	126	126	126
Language Families	54	54	54	28	28	28

Notes: This table explores the relative contributions of ecological diversity in the contemporary homeland vs. the Urheimat to the presence of politeness distinctions in a daughter language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S9.8: Geographic Origins of Jurisdictional Hierarchy and Politeness Distinctions

	Mechanism				
	Jurisdictional Hierarchy		Politeness Distinctions		
	OLS	OLS	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)
Ecological Diversity	0.24*** (0.06)	0.12*** (0.04)			
Average Caloric Yield (All Crops, pre-1500)	0.18*** (0.07)	0.24*** (0.04)			
Jurisdictional Hierarchy			0.11*** (0.02)	0.10*** (0.02)	0.15*** (0.05)
All Geographic Controls	No	Yes	No	Yes	Yes
Regional FE	No	Yes	No	Yes	Yes
First-stage F-statistic					14.24
Adjusted- R^2	0.07	0.32	0.18	0.41	0.38
Observations	1137	1137	139	139	139

Notes: This table documents the positive statistically and economically significant association between the geographical determinants of statehood, as measured by jurisdictional hierarchy beyond the local level, and politeness distinctions in a language. The first two columns provide the evidence on the association between ecological diversity and statehood, and columns (3) and (4) show the association between statehood and the emergence of politeness distinctions. The table shows the estimated coefficients in an OLS regression as the dependent variable in columns (1) and (2) is not binary. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S9.9: Hierarchy and Politeness (Alternate Measure)

	Hierarchy Index			
	All		Old World	
	(1)	(2)	(3)	(4)
Politeness Distinctions	0.31*** (0.07)	0.37*** (0.11)	0.36*** (0.09)	0.43*** (0.12)
Ecological Diversity		0.18* (0.10)		0.08 (0.08)
Regional FE	Yes	Yes	Yes	Yes
Main Geographical Controls	No	Yes	No	Yes
Adjusted- R^2	0.42	0.49	0.42	0.57
Observations	53	53	50	50

Notes: This table documents the statistically and economically positive association between politeness distinctions and preferences for hierarchy in society. The analysis accounts for the geographical origins of politeness distinctions and other geographical characteristics of the contemporary homeland of the language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

S9.B Alternative Measure of Politeness Distinctions (Intensive Margin)

Table S9.10: Geographic Origins of Politeness Distinctions (Intensive measure - OLS)

	Intensity of Politeness Distinctions					
	(1)	(2)	(3)	(4)	(5)	(6)
Ecological Diversity	0.26*** (0.06)	0.21*** (0.07)	0.20*** (0.07)	0.21*** (0.08)	0.21*** (0.08)	0.22*** (0.07)
All Crops (Average Caloric Yield, pre-1500)		0.14*** (0.05)	0.14*** (0.05)	0.14*** (0.05)	0.14*** (0.05)	0.19*** (0.06)
Absolute Latitude		0.02 (0.05)	0.02 (0.06)	0.16 (0.16)	0.17 (0.15)	-0.01 (0.15)
Elevation		-0.08 (0.07)	-0.06 (0.07)	-0.02 (0.09)	-0.01 (0.08)	0.01 (0.08)
Ruggedness		-0.01 (0.05)	-0.02 (0.06)	-0.01 (0.06)	-0.00 (0.06)	-0.06 (0.06)
Coast Length		0.12 (0.09)	0.11 (0.09)	0.12 (0.10)	0.12 (0.10)	0.10 (0.09)
Precipitation (mm/month)			-0.01 (0.08)	-0.02 (0.10)	-0.01 (0.09)	-0.10 (0.10)
Precipitation (mm/month) (std)			0.12 (0.08)	0.22 (0.14)	0.23 (0.14)	0.26** (0.13)
Precipitation Volatility			-0.06 (0.09)	-0.10 (0.10)	-0.10 (0.09)	-0.06 (0.10)
Precipitation Spatial Correlation			-0.03 (0.06)	0.45 (0.44)	0.46 (0.45)	0.41 (0.45)
Temperature (Daily Mean)				0.05 (0.14)	0.04 (0.13)	-0.03 (0.13)
Temperature (Daily Mean) (std)				-0.13 (0.12)	-0.15 (0.12)	-0.23** (0.11)
Temperature Volatility				-0.10 (0.12)	-0.06 (0.12)	0.11 (0.12)
Temperature Spatial Correlation				-0.51 (0.45)	-0.52 (0.46)	-0.44 (0.46)
Unproductive Period (pre-1500CE)					-0.10** (0.05)	-0.12** (0.05)
Regional FE	No	No	No	No	No	Yes
Adjusted- R^2	0.11	0.17	0.17	0.17	0.18	0.32
Observations	198	198	198	198	198	198

Notes: This table documents the positive, statistically, and economically significant association between ecological diversity and an intensive measure of politeness distinctions in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S7.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S9.11: Geographical Origins of and Politeness Distinctions (Intensive Measure - OLS)
Robustness to Spatial-Autocorrelation, Clustering and Selection

	Intensity of Politeness Distinctions					
	(1)	(2)	(3)	(4)	(5)	(6)
Ecological Diversity	0.26*** (0.06) ([0.07]) [0.07] {0.05}	0.21*** (0.07) ([0.09]) [0.09] {0.05}	0.20*** (0.07) ([0.09]) [0.09] {0.05}	0.21*** (0.08) ([0.09]) [0.09] {0.06}	0.21*** (0.08) ([0.09]) [0.09] {0.06}	0.22*** (0.07) ([0.07]) [0.08] {0.05}
Geographical Controls	No	Yes	Yes	Yes	Yes	Yes
Regional FE	No	No	No	No	No	Yes
Altonji et al						28.16
δ						49.64
β -Oster						0.22
R^2	0.11	0.19	0.21	0.23	0.25	0.39
Observations	198	198	198	198	198	198

Notes: This table documents the positive, statistically, and economically significant association between ecological diversity and the existence of politeness distinctions in the language spoken in this region, accounting for regional fixed-effects and other geographical characteristics. All columns have the same specification and controls as Table S4.1. Geographical controls include absolute latitude, mean elevation, terrain ruggedness, and coast length, as well as other agriculture-related controls as precipitation and temperature means and standard deviations. All independent variables have been normalized by subtracting their mean and dividing by their standard deviation. Thus, all coefficients can be compared and show the impact of a one standard deviation in the independent variable on the probability of having a future tense in the language. Heteroskedasticity robust standard error estimates are reported in parentheses, clustered at the language genus in parenthesis and squared brackets, spatial auto-correlation corrected standard errors (*101*) in squared brackets and Cliff-Ord ML in curly brackets; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S9.12: Persistent Impact of Geographical Characteristics on Politeness (Intensive Measure)
Linguistic Homeland vs. Urheimat
(Languages Outside Urheimat)

	Intensity of Politeness Distinctions					
	Migratory Distance to Urheimat					
	Any Distance			At Least 1 Week		
	Homeland	Urheimat	Both	Homeland	Urheimat	Both
	(1)	(2)	(3)	(4)	(5)	(6)
Homeland's Ecological Diversity	0.26*** (0.08)		0.21** (0.08)	0.28*** (0.08)		0.26*** (0.07)
Urheimat's Ecological Diversity		0.03 (0.07)	-0.05 (0.08)		0.17 (0.11)	-0.03 (0.09)
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes
Homeland Geographical Characteristics	Yes	No	Yes	Yes	No	Yes
Urheimat Geographical Characteristics	No	Yes	Yes	No	Yes	Yes
Adjusted- R^2	0.34	0.29	0.42	0.39	0.39	0.57
Observations	180	180	180	126	126	126
Language Families	54	54	54	28	28	28

Notes: This table explores the relative contributions of ecological diversity in the contemporary homeland vs. the Urheimat to the presence of politeness distinctions in a daughter language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S9.13: Geographic Origins of Jurisdictional Hierarchy and Politeness Distinctions
(Intensive Measure)

	Mechanism				
	Jurisdictional Hierarchy		Politeness Distinctions		
	OLS	OLS	OLS	OLS	IV
	(1)	(2)	(3)	(4)	(5)
Ecological Diversity	0.24*** (0.06)	0.12*** (0.04)			
Average Caloric Yield (All Crops, pre-1500)	0.18*** (0.07)	0.24*** (0.04)			
Jurisdictional Hierarchy			0.38*** (0.05)	0.33*** (0.06)	0.57*** (0.11)
All Geographic Controls	No	Yes	No	Yes	Yes
Regional FE	No	Yes	No	Yes	Yes
First-stage F-statistic					14.24
Adjusted- R^2	0.07	0.32	0.35	0.54	0.43
Observations	1137	1137	139	139	139

Notes: This table documents the positive statistically and economically significant association between the geographical determinants of statehood, as measured by jurisdictional hierarchy beyond the local level, and politeness distinctions in a language. The first two columns provide the evidence on the association between ecological diversity and statehood, and columns (3) and (4) show the association between statehood and the emergence of politeness distinctions. The table shows the estimated coefficients in an OLS regression as the dependent variable in columns (1) and (2) is not binary. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S9.14: Hierarchy and Politeness

	Hierarchy Index					
	All			Old World		
	(1)	(2)	(3)	(4)	(5)	(6)
Politeness Distinctions	0.48*** (0.15)	0.37** (0.14)	0.37** (0.16)	0.55*** (0.16)	0.56*** (0.16)	0.56*** (0.10)
Ecological Diversity			0.18* (0.09)			0.07 (0.09)
Regional FE	No	Yes	Yes	No	Yes	Yes
Main Geographical Controls	No	No	Yes	No	No	Yes
Adjusted- R^2	0.22	0.43	0.49	0.28	0.47	0.60
Observations	53	53	53	50	50	50

Notes: This table documents the statistically and economically positive association between of politeness distinctions and preferences for hierarchy in society. The analysis accounts for the geographical origins of politeness distinctions and other geographical characteristics of the contemporary homeland of the language. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

S10 Variable Definitions, Sources and Summary Statistics

S10.A Variable Definition and Sources

- **Absolute latitude:** The absolute value of the latitude of a homeland's approximate geodesic centroid. Author's computations.
- **Mean Elevation:** The mean elevation of a homeland in km above sea level, calculated using geospatial elevation data taken from (102). Author's computations.
- **Terrain Ruggedness:** The mean change in elevation across cells in a homeland in km, calculated following the methodology of (103), using geospatial elevation data taken from (102). Author's computations.
- **Caloric Suitability:** Pre-1500CE Caloric suitability is the potential caloric output in a region as reported in (4).
- **Coast length:** Length, in thousands of km, of a country's coastline. Author's computations.
- **Ecological Diversity:** Herfindahl index of share's of a country's area in various ecologies. Author's computations following the method of (18) and (97).
- **Volatility (temperature and precipitation):** Volatility of temperature and precipitation constructed using v3.2 of the Climatic Research Unit (CRU) database following the method of (104).
- **Diversification (temperature and precipitation):** Spatial Correlation of temperature and precipitation shocks constructed using v3.2 of the Climatic Research Unit (CRU) database following the method of (104).

S10.B Summary Statistics

Table S10.1: Summary Statistics of the Existence of Periphrastic Future Tense by Region

Region	Observations	Mean	Std. Dev.
Sub-Saharan Africa	66	0.53	0.503
Middle East and North Africa	8	0.5	0.53
Europe and Central Asia	56	0.48	0.50
South Asia	21	0.19	0.40
East Asia and Pacific	71	0.55	0.50
North America	22	0.41	0.50
Latin America	31	0.55	0.50
Total	275	0.49	0.50

Table S10.2: Summary Statistics of the Existence of Sex-Based Grammatical Gender Systems by Region

Region	Observations	Mean	Std. Dev.
Sub-Saharan Africa	27	0.63	0.49
Middle East and North Africa	7	0.71	0.49
Europe and Central Asia	40	0.48	0.51
South Asia	16	0.63	0.50
East Asia and Pacific	70	0.27	0.45
North America	25	0.08	0.28
Latin America	32	0.28	0.46
Total	227	0.37	0.48

Table S10.3: Summary Statistics of the Existence of Politeness Distinctions by Region

Region	Observations	Mean	Std. Dev.
Sub-Saharan Africa	36	0.14	0.35
Middle East and North Africa	4	0.25	0.50
Europe and Central Asia	34	0.71	0.46
South Asia	19	0.63	0.50
East Asia and Pacific	59	0.32	0.47
North America	18	0.00	0.00
Latin America	28	0.18	0.39
Total	207	0.34	0.48

Table S10.4: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Absolute Latitude	0.096	1.025	-1.302	2.613	275
Elevation	0.027	1.026	-0.92	4.827	275
Ruggedness	-0.014	0.979	-0.877	6.162	275
Coast Length	0.024	1.154	-0.302	11.692	275
Precipitation	-0.078	0.928	-1.3	4.4	275
Precipitation (std)	-0.02	0.911	-0.667	8.314	275
Precipitation Volatility	-0.064	0.926	-1.531	4.665	275
Precipitation Spatial Correlation	0.064	0.939	-2.133	0.810	275
Temperature (Daily Mean)	-0.054	0.977	-2.996	1.176	275
Temperature (Daily Mean) (std)	-0.017	0.929	-0.877	4.876	275
Temperature Volatility	0.079	0.991	-1.641	3.504	275
Temperature Spatial Correlation	0.068	0.939	-2.161	0.683	275

Table S10.5: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Intensity of Agriculture	8.890	3.061	2	12	264

Figure S11.1: Location of Urheimats for 147 Language Families in Ethnologue

S11 Identifying the Location of Urheimats of Language Families

The location of language's Urheimat (i.e., the geographical region where the language family emerged) is identified based on one methodology proposed by linguists (23), according to which the geographical area that maximized linguistic diversity within a language family corresponds to the Urheimat.⁴⁸

In particular, using data on the linguistic diversity of 4169 languages from version 12 of the ASJP (105), (23) identify the Urheimat of 82 language families based on the comparison of the level of linguistic diversity of each language within each language family. Following their methodology the analysis identifies the location of Urheimats for 147 language families based on the updated version 18 of the ASJP (106). Figure S11.1 depicts the location of these Urheimats, while Figures S11.2 and S11.3 depict the extent of the Indo-European and Niger-Congo language families and the location of their Urheimats.

⁴⁸This method builds upon a similar methodology employed by geneticists to determine the location in which domestication of crops was originated.

Figure S11.2: Indo-European Language Family: Homelands and Urheimat

Figure S11.3: Niger-Congo Language Family: Homelands and Urheimat

S11.A Robustness Analysis - Alternate Urheimat Location

Given the uncertainty surrounding the identification of the location of the Urheimat, this section explores the robustness of the results in Table 2 to the location of the Urheimat. Specifically, using the list of sources from the historical linguistics literature presented in Table S11.1, the analysis locates alternative approximate regions where Urheimats for the language families in the analysis are hypothesized to be located. Given the limited information on some language families, this procedure provides a smaller sample of language families for which alternative sources were identified than the sample of language families in the main text. Since the exact location is less precisely determined in this method there is some additional measurement error and uncertainty in the estimation process. This is especially true for the case for the measure of ecological diversity, which depends strongly on the location and shape of the Urheimat. Specifically, for each geographical measure, the analysis computes the average of the among the languages located in the proximity of the hypothesized location of the Urheimat. Given the fuzzy nature of the location the analysis cannot identify migratory distances from the Urheimat, preventing the full replication of the results in the main text. Figure S11.4 depicts the correlation between the measures of the main independent variables in Table 2, and shows that the measures of caloric crop return and caloric plow suitability are very highly correlated with the main measure employed in the main text. Moreover, replicating the analysis of Table 2 with this data generates similar results, albeit less statistically significant (Table S11.2). As expected, the measures of ecological diversity are much less correlated, and the results on politeness are less statistically significant.

Figure S11.4: Correlation between main independent variables across Urheimat sources

Table S11.1: Language Families, Languages in the Sample and Sources for Urheimat

Language Family	Languages in the Family	Sources
Afro-Asiatic	Afar, Alagwa, Amharic, Egyptian Arabic, Gulf Speaking Arabic, Tunisian Arabic, Arbore, Awngi, Baiso, Bedawiyet, Dera, Dizi, Hausa, Hebrew, Iraqw, Kera, Kistane, Lele, Libido, Maltese, Central Marghi, Miya, Eastern Oromo, Pero, Rendill, Sebat Bet Gurage, Sidamo, Somali, Tamhaq, Tahggart, Central Atlas Tamazight, Tigre, Tigrigna, Wolaytta	(107)
Ainu	Orok	(108)
Algic	Cheyenne, Cree: Plains, Montagnais, Yurok	(23), (109)
Altaic	South Azerbaijani, Buriat: Mongolia, Chuvash, Evenki, Karaim, Kyrgyz, Mongolian Halh, Khalaj: Turkic, Orok, Turkish, Uyghur, Uzbek: Northern, Yakut	(110)
Arauan	Paumari	(111)
Araucanian	Mapudungun	(112)
Arawakan	Apurina, Garifuna, Ashaninka	(113)
Australian	Alawa, Alyawarr, Bandjalang, Dyirbal, Gooniyandi, Guguymidjir, Gunwinggu, Jaru, Kayardild, Koyata, Kuku-Yalangi, Mangarayi, Martu Wangka, Martuyhunira, Maung, Nangikurrunggur, Nungubuyu, Pitjantjatara, Tiwi, Warjarri, Wambaya, Wangaaybuwan-Ngiyambaa, Worora, Yidinyi, Yindjibarndi	(114)

Continued on next page

Table S11.1 – continued from previous page

Language Family	Languages in the Family	Sources
Austro-Asiatic	Khazi, Central Khmer, Khmu, Nicobarese: Car Palaung: Ruching, Semelai, Vietnamese	(115)
Austronesian	Batak Karo, Batak Toba, Buli, Cebuano, Chamorro, Chuukese, Fijian, Halia, Hawaiian, Iai, Javanese, Kambera, Kilivila, Kosraean, Makian: East, Malagasy: Plateu, Malay, Maori, Mbula, Mokilese, Motu, Nakanai, Paiwan, Palauan, Pangasinan, Pohnpeian, Rapa Nui, Rukai, Sabu, Sie, Sinaigoro, Tuvaluan, Sunda, Tagalog, Tahitian, Tangga, Tawala, Tetun, Tukang Besi South, Ura, Uripiv-Wala-Rano-Atchin, Vaekau-Taumako	(116)
Aymaran	Aymara: Southern, Jaqaru	(113)
Baining-Taulil	Qaget	(117)
Barbacoan	Awa-Cuaiquer	(112)
Basque	Basque	(22)
Border	Imonda	(118)
Burushaski	Burushaski	(22)
Caddoan	Wichita	(23), (109)
Cariban	Carib, Hixkaryana	(119)
Chapacura-Wanham	Pakaasnovos	(112)
Chibchan	Arhuaco, Bribri, Ngabere, Rama	(113)
Choco	Epena	(119)
Chukotko-Kamchatkan	Chukchi, Itelmen	(23, 113)

Continued on next page

Table S11.1 – continued from previous page

Language Family	Languages in the Family	Sources
Dagan	Daga	(120)
Dravidian	Brahui, Kolami: Northwestern, Kannada, Kui, Malayalam, Tamil, Telugu	(121)
Eskimo-Aleut	Aleut, Inuktitut: Eastern Canadian, Inuktitut: Greenlandic, Yupik: Central	(122), (23)
Gapun	Taiap	(123)
Guaicuruan	Mokovi	(112)
Haida	Haida: Northern	(124)
Hmong-Mien	Hmong-Djua	(23, 125)
Hokan	Cocopa, Maricopa, Pomo: Central, Pomo: Southeastern	(23) (109)
Indo-European	Afrikaans, Albanian: Gheg, Armenian, Eastern Balochi, Belarusian, Bengali, Bosnian, Breton, Bulgarian, Catalan-Valencian-Balear, Croatian, Czech, Danish, Dutch, English, Western Farsi, French, Western Frisian, Greek, Gujarati, Hindi, Icelandic, Italian, Kashmiri, Gaelic: Irish, Gaelic: Scottish Gaelic, Standard German, Northern Kurdish, Latvian, Lithuanian, Luxembourgish, Macedonian, Maithili, Marathi, Nepali, Norwegian, Ossetian, Eastern Panjabi, Pashto: Central, Polish, Portuguese, Romanian, Romansh, Russian, Savi, Seraiki, Slovak, Slovene, Spanish, Swedish, Ukrainian, Urdu, Serbian, Sindhi, Sinhala, Vlaams, Welsh, Yazgulyam	(126)

Continued on next page

Table S11.1 – continued from previous page

Language Family	Languages in the Family	Sources
Iroquoian	Oneida, Seneca	(23), (109)
Japanese	Japanese	(127)
Jivaroan	Shujar	(111)
Kadugli	Krongo	(23)
Karok	Karok	(109)
Kartvelian	Georgian	(23)
Keresan	Western Keres	(109)
Khosian	Juhoan, Nama	(22)
Kiowa-Tanoan	Kiowa	(109)
Korean	Korean	(22)
Kutenai	Kutenai	(109)
Lower Sepik-Ramu	Yimas	(23)
Macro-Ge	Canela	(128)
Marind	Marind	(129)
Matacoan	Wichi Lhaments Guisnay	(119)
Mayan	Eastern Jakalteko, Maya: Yucatan, Tojolabal	(113)
Mixe-Zoque	Zoque: Copainala	(113)
Mosetenan	Tsimane	(130)
Mura	Piraha	(112)
Muskogean	Koasati	(109)
Na-Dene	Slavey: North, Tlingit	(23), (109)
Nakh-Daghestanian	Archi, Dido, Ghodoberi, Hunzib, Ingush, Lak, Lezgi	(113)

Continued on next page

Table S11.1 – continued from previous page

Language Family	Languages in the Family	Sources
Niger-Congo	Akan, Bamanakan, Baoule, Bemba, Dagbani, Defaka, Engenni, Ewe, Ganda, Gikuyu, Southern Grebo, Igbo, Isekiri, Jola-Fonyi, Western Karaboro, Koongo, Koromfe, Lelemi, Lozi, Luvale, Mann, Kissi: Northern, Moore, Mwera, Noon, Nyanja, Pular, Rwanda, Senufo: Supiyre, Northern Sotho, Southern Sotho, Swahili, Swati, Tem, Themne, Tooro, Tsonga, Tswana, Venda, Vengo, Wolof, Xhosa, Yoruvba, Zanda, Zulu	(43)
Nilo-Saharan	Avokaya, Bagirmi, Bari, Didinga, Fur, Kenuzi-Dongola, Central Ksnuri, Kunama, Lango, Masalit, Murle, Ngambay, Ngiti, Songhay: Koyraboro Senni	(132)
Nimboran	Nimboran	(123)
Nivkh	Gilyak	(133)
Northwest Caucasian	Abkhaz	(23)
Oregon Coast	Coos	(111)
Oto-Manguean	Chinantec: Palantla, Mixtec: San Miguel el grande, Otomi: Mezquital	(134)
Panoan	Chacobo, Shipibo-Conibo	(119)
Peba-Yaguan	Yagua	(135)
Penutian	Maidu: Northeast, Nez-Perce, Tsimshian, Umatilla	(23), (109)
Quechuan	Quechuan: South Bolivian, Quichua: Imbabura Highland	(111)
Salishan	Shuswap, Thompson	(23), (109)

Continued on next page

Table S11.1 – continued from previous page

Language Family	Languages in the Family	Sources
Sentani	Sentani	(23)
Sepik	Alamblak, Yessan-Mayo	(123)
Sino-Tibetan	Burmese, Chepang, Chin: Bawm, Chin: Hakk, Chinese: Hakka, Chinese: Mandarine, Chinese: Yue, Dumi, Garo, Kayah: Eastern, Ladakhi, Lahu, Lepcha, Limbu, Meiti, Rawang, Tibetan: Central, Wayu	(136)
Siouan	Lakota	(23), (109)
Solomons-East Papuan	Lavukaleve	(137)
Tacanan	Araona	(119)
Tai-Kadai Tol	Dong: Northern, Lao, Thai Tol	(138)
Torricelli	Bukiyip	(23)
Totonacan	Totonac: Yecuatla	(119)
Trans-New Guinea	Amele, Bongu, Citak: Tamnim, Dani: Lower Grand Valley, Kewa: West, Kobo, Kombai, Ono, Suena, Una, Usan, Yagaria	(23)
Tucanoan	Barasana-Eduria, Tanimuca: Retu-ara, Tucano	(112)
Tupian	Cocama-Cocamilla, Siriono, Udmurt	(139)
Uralic	Erzya, Finnish, Hungarian, Kaapor, Mansi Nenets,	(140)
Uto-Aztecan	Cahuilla, Comanche, Nahuatl: Eastern Huasteca, Nahuatl: Northern Puebla, Nahuatl: Telelcingo, Pipil, Timbisha, Tohono-Oodham, Yaqui	(109), (23)

Continued on next page

Table S11.1 – continued from previous page

Language Family	Languages in the Family	Sources
Wakashan	Nootka	(23), (109)
Warao	Warao	(141)
West Papuan	Abun, Hatam, Mai Brat, Tidore	(23)
Yanomam	Sanuma	(141)
Yele	Yele	(142)
Yeniseian	Ket	(119)
Yukaghir	Southern Yukaghir	(143)
Zuni	Zuni	(109)
Other	Angolar, Aukan, Bunaba, Hawaiian Creole English, Tayo	

S12 The Timing of the Appearance of Ancient Plows

The hypothesized coevolution of gender biases and sex-based grammatical gender based on their common geographical origin necessitates the appearance of the plow prior to the emergence of significant fraction of daughter languages within each language family. Indeed, paleo-linguistic and archeological evidence suggest that the plow appears prior to the formation of languages with a sex-based grammatical gender. Evidence indicates that the plow has originated during the emergence of agriculture. In particular, evidence suggests that sporadic use of the plow is present as early as 10-13 thousands years ago in the fertile crescent, and in particular along the Tigris, Euphrates, Nile, as well as in Indus and Yangtze River valleys (144). Primitive prototypes of plows that were sufficiently demanding physically to provide a comparative advantage in their use for male (e.g, the *ard*) were more prevalent in the fertile crescent and in particular in Mesopotamia and along the Nile river 5000-8000 years ago (144, 145). Moreover, nearly 4000 years ago, the plow was adopted by the Greeks, was further developed by the Indus Valley civilizations (i.e., the forefathers of the Dravidian language family), and by civilizations that resided along the Yangtze river (145). In addition, the plow was prevalent in both Northern Africa and Ethiopia, where the ancestors of the Afro-Asiatic language family resided, at least 2500 years ago (146).

Linguistic evidence suggests that early Indo-Europeans used the plow during the late Neolithic period, before their proto-language separated into its different daughter languages around

Table S11.2: Persistent Effect of Urheimat Characteristics
Languages Outside Urheimat (Alternative Urheimat)

	Existence of Language Structure		
	Gender	Politeness	Future
	(1)	(2)	(3)
Homeland's Caloric Plow Suitability	-0.06 (0.07)		
Urheimat's Caloric Plow Suitability	0.14* (0.07)		
Homeland's Ecodiversity		0.07 (0.05)	
Urheimat's Ecological Diversity		0.35 (0.28)	
Homeland's Caloric Crop Return			-0.03 (0.06)
Urheimat's Caloric Crop Return			0.41*** (0.07)
Regional FE	Yes	Yes	Yes
Homeland Geographical Characteristics	Yes	Yes	Yes
Urheimat Geographical Characteristics	Yes	Yes	Yes
Adjusted- R^2	0.72	0.47	0.21
Observations	100	116	163
Language Families	20	19	19

Notes: Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests; All regressions include a constant.

3000 BCE. In particular, German emerged around the 1st century BCE, English and Spanish in about the 5th century CE and Italian around the 14th century CE. The proto-language spoken around 4500 BCE consists of a word for the plow prior to the demic diffusion of the Indo-European languages to Asia and Europe (147). Reconstructions of the proto-Indo-European language suggest that the verb, plow, was *ar(e)*, which is the root of *ear* (English), *erian* (Old English), *erja* (Old Norse), *arjan* (Gothic), *arare* (Latin), *arar* (Spanish), as well as *ard* (English) and *arado* (Spanish) (148).

In contrast to temperate zones where plow positive crops were suitable for cultivation, tropical zones as well as the Americas have not adopted the earlier versions of the plow. Instead, in order to till the soil these areas adopted a hoe (i.e., a pointy stick used by hunter-gatherers) that

Figure S12.1: Ancient Egyptian ard. Burial chamber of Sennedjem (ca. 1200 BCE).

Figure S12.2: Gang of men plowing and sowing. Babylonian Seal Impression (1500 BCE)

was more suitable for their climatic environment. In these areas, the plow was introduced by

Figure S12.3: Gang of men plowing and sowing. Sumerian (ca.4500-2000BCE).

Figure S12.4: Picture of a field plowed along the Indus river (ca. 2000 BCE)

Europeans during the colonial period.⁴⁹

⁴⁹Figures S12.1-S12.3 depict ancient plow agriculture in Egypt, Babylon and Sumer around 2000BCE. Additionally, Figure S12.4 depicts a plowed field in the Indus Valley around 2000BCE. Literary evidence for the early adoption of the plow in these regions is “Debate between the pickaxe and the plough” (*The Sumerian Disputation/Debate*; 3rd millennium BCE), the Old and the New Testaments, and the Babylonian *Code of Hammurabi* (ca. 1750BCE) where references to the plow or plowing is present in laws 43, 44, 242, and 260.

S13 Supplemental Tables

Table S13.1: Geographic Origins of Mechanism and Language Structure

	Mechanism			Language Structure		
	Agr.Intensity	Plow	Hierarchy	Future	Gender	Politeness
	(1)	(2)	(3)	(4)	(5)	(6)
Caloric Crop Return	0.24*** (0.03)					
Caloric Plow Suitability		0.06** (0.02)				
Ecological Diversity			0.12*** (0.03)			
Agricultural Intensity				0.08* (0.04)		
Aboriginal Plow					0.23** (0.11)	
Jurisdictional Hierarchy						0.16*** (0.03)
All Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted- R^2	0.64	0.47	0.32	0.17	0.31	0.48
Observations	1303	1175	1154	263	145	139

Notes: This table documents, based on OLS regressions, the statistically and economically significant association between each geographical determinant and the proposed mechanism (columns 1-3). Additionally, it establishes, based on OLS regressions, the statistically and economically significant association between each proposed mechanism and the corresponding language structure (columns 4-6). Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S13.2: Geographic Origins of Mechanism and Language Structure
Instrumental Variable Regression

	Mechanism (First-Stage)			Language Structure (IV)		
	Agr.Intensity	Plow	Hierarchy	Future	Gender	Politeness
	(1)	(2)	(3)	(4)	(5)	(6)
Crop Return (pre-1500CE)	0.24*** (0.03)					
Caloric Plow Suitability		0.06** (0.02)				
Ecological Diversity			0.12*** (0.03)			
Intensity of Agriculture				0.21* (0.11)		
Aboriginal Plow					0.95*** (0.31)	
Jurisdictional Hierarchy						0.32*** (0.06)
All Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F-statistic				29.13	15.96	16.15
Adjusted- R^2	0.64	0.47	0.32	0.13	0.06	0.35
Observations	1303	1175	1154	263	145	139

Notes: This table documents, based on instrumental variable regressions, the statistically and economically significant association between each geographical determinant and the proposed mechanism (columns 1-3, first stage). Additionally, columns 4-6 show the second stage and establish the statistically and economically significant association between each proposed mechanism and the corresponding language structure (columns 4-6). Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

S14 Uniqueness of Associations and Multiple Hypothesis Testing

This section explores the uniqueness of the associations between geographical origins and linguistic traits and expands on the findings shown in Figure 3. As is clarified in sections S1.A.1 (for long-term orientation and the periphrastic future tense), S1.A.2 (for gender bias and sex-based grammatical gender) and S1.A.3 (for hierarchy and politeness distinctions), the analysis is guided by socio-linguistic hypotheses that linked these important cultural characteristics to linguistic traits that could have reinforced them. Nonetheless, the associations between geographical and linguistic traits may reflect a broader association between each of these geographical factors and a wide range of linguistic traits. Moreover, one may be concerned about the need to correct for multiple hypotheses testing since an atheoretical approach to the question could have instead searched systematically for any underlying associations (in the spirit of genome wide association studies–GWAS). To address these issues, this section (i) explores how the underlying geographical determinants may be associated with other linguistic traits, and (ii) corrects for multiple hypothesis testing.

Tables S14.1-S14.9 establish that each of the geographical origins is uniquely associated with their complementary linguistic traits. Indeed, these tables suggest that alternative language structures (as reported by WALS), such as the existence of other temporal and non-temporal structures, for which there do not seem to exist socio-linguistic theories linking them to the the studied cultural values, are largely uncorrelated with the proposed geographical determinant. For example, pre-1500 crop return is not significantly associated with the existence of either the past or the perfect tense, and it is orthogonal to non-temporal linguistic characteristics such as: (i) the existence of possessive classifications, (ii) the existence of coding for evidentiality, (iii) the number of consonants, and (iv) the number of colors. Moreover, as documented in Table S14.2, pre-1500 crop return is not associated with any of the temporal structures analyzed in WALS. See sections S4-S6 for further discussion of each specific relation.

In order to further explore the uniqueness of the associations in the main analysis, as well as the importance of the origins, we explore the semi-partial R^2 of the hypothesized origins relative to the sum of the semi-partial R^2 of all other variables included in the analysis for their hypothesized complementary linguistic traits, as well as all other linguistic traits available in WALS. Specifically, the semi-partial R^2 of a variable measures the share of the total variation in the outcome variable that is *uniquely* associated with it, i.e., after partialling out the effect of all other variables included in the regression analysis on the proposed origin. Figure S14.1 depicts the semi-partial R^2 for the hypothesized geographical origins for their complementary traits as well as all for all other linguistic traits in WALS. As can be seen, the semi-partial R^2 of the proposed geographical origins in the association with their complementary linguistic traits are much larger (4-8 times the semi-partial with other traits in WALS). Moreover, this figure and Figure S14.2 establish that the variation uniquely associated to these hypothesized geographical origins explains more of the variation of their complementary traits than of other variables in WALS, relative to other linguistic traits in WALS. Specifically, Figure S14.2 shows

the ratio of the semi-partial R^2 of the hypothesized origin to the sum of the semi-partial R^2 of all other variables in the complementary linguistic traits and with other traits in WALS. The ratio is significantly larger for the complementary traits than for other traits in WALS.

To explore the uniqueness of the hypothesized relations further, Figures S14.3-S14.5 depict the estimated coefficients, and the distribution of p-values, t-statistics and estimates when estimating the main specification using alternative measures from WALS as outcome variables. As in the previous analysis, this is a very conservative test, as it does not take into consideration that (i) the presence of linguistic traits may correlate across languages (i.e., any language that has linguistic trait X also has linguistic trait Y), which may create statistically significant associations to geographical factors, but not through the hypothesized cultural channel. (ii) Since the analysis holds the set of controls constant across specifications, statistically significant associations may reflect omitted variables or other causal links between the mediating mechanisms identified in this paper and other linguistic traits. E.g., the analysis has established that crop return is associated with the adoption of agriculture, which itself is associated with the emergence of hierarchical structures. Thus, it is unsurprising that crop return is also associated with the prevalence of politeness distinctions. (iii) Differences in sample sizes due to data availability in WALS, which may allow some associations to be estimated more precisely than others. And, (iv) the coding of these WALS measures may not be sensible, creating false statistically significant associations when there are none. E.g., in WALS the linguistic trait “144D The Position of Negative Morphemes in SVO Languages” is coded as follows: 1 NegSVO, 10 SOV/SOV but SNegVO, 11 Other NegV, 13 More than one construction, 14 ObligNeg, 15 OptNeg, 2 SNegVO, 4 SVONeg, 5 S[Neg-V]O, 6 S[V-Neg]O, 7 SVO&Tone, 12 Other VNeg, 3 SVNegO, 8 VSO but NegSVO, 9 SVO/SOV but SVONeg. Clearly, these quantitative values do not have any real quantitative meaning (in fact different statistical software may convert these to different categorical variables and thus assign even more different values and orderings). Thus, these tests will tend to find associations when none are present, i.e. generate false positives or Type I statistical errors, which will penalize excessively the estimation of our main results by affecting the distribution of all estimated p-values. Nonetheless, even under these very conservative tests, the analysis finds very few significant associations between the proposed geographical origins and other linguistic traits.

To appreciate this better, Figure S14.6 depicts the logarithm of p -values against the logarithm of the absolute value of the estimated effect of the geographical origin for all these regressions. It is noticeable that almost none of the other potential associations have similarly large effect sizes and low p -values. To further explore these relations, Figure S14.7 replicates the analysis on the set of variables that are strictly binary or can be directly converted to a binary coding. E.g., WALS trait 16A “Weight Factors in Weight-Sensitive Stress Systems” can be coded into a binary variable reflecting the presence of weight factors. Of course there are variables that cannot be converted into a binary coding (as the example of trait 144D above shows). The results are qualitatively similar and suggest that the hypothesized geographical roots are indeed uniquely associated with their linguistic traits. As a final test, the analysis was performed after encoding all variables into a binary coding. So, e.g., linguistic trait 144D was recoded as

15 binary variables, each reflecting the presence of each of the categories in the original WALS variable. Figure S14.8 depicts the results of this analysis. Reassuringly, the qualitative results remain unchanged. Finally, Figure S14.9 replicates the analysis using all binary coded WALS variables on the sample of languages in which all three main linguistic features are available. Again the qualitative results remain unchanged. In fact, the hypothesized relations tend to be among the 5% smallest p -values and the largest 5% estimated effects across all these various analyses. These analyses suggest that the hypothesized relations between each geographical root and its associated linguistic trait are indeed unique in terms of statistical significance and effect size.

In the results of these two previous analyses a few linguistic traits that had not been identified as complementary to the cultural traits have very strong associations with some of the geographical origins. Some of these seem to reflect causal links between the proposed mechanisms and other structures. E.g., politeness distinctions are strongly associated with crop return, reflecting the association of crop return with the adoption of agriculture and the effect of agriculture on the emergence of states and hierarchical structures (43, 97). Interestingly, some seem to reflect other complementary linguistic traits that had not been identified by us when exploring the socio-linguistic literature, but for which there does exist socio-linguistic theories supporting their complementary role. E.g., the presence of passive constructions (variable 107A in WALS) has a very strong association with ecological diversity, which seems to reflect its role in politeness (5, 99). This can be seen as an out of sample prediction that further supports the hypothesis of this paper.

The analysis additionally explores the robustness of the results to multiple hypothesis testing. Clearly, as the previous analyses show, one could in principle perform similar tests using any of the variables available in WALS. Table S14.10 corrects for false discovery rates following the procedures of (149) and (150) using all the data in WALS. The analysis in Table S14.10 replicates the main specification for each geographical origin on all original WALS data. As explained above, these analyses are conservative and work against the main hypothesis, as they are prone to generate false positives due to e.g., inclusion of variables which have quantitative codings that may not make any sense, thus affecting the distribution of p -values by generating more erroneously low p -values. The results shown in this table suggest that even under these strong assumptions, the association between the hypothesized geographical origins and these linguistic traits are statistically significant at conventional values. Additionally, replicating the analysis for only the binary coded variables increases the statistical significance of the results as shown in Table S14.11. Tables S14.12-S14.23 explore the effect of correcting for false discovery rates under various assumptions including clustering at the family level or not, different ways of coding WALS variables, and selecting common samples for the variables. As can be seen the results do depend on the exact assumptions used for multiple hypothesis testing. Nonetheless, even under these very conservative estimates the results remain robust at traditional levels of confidence.

Table S14.1: Geographic Origins of Future and Language Structures

	Language Structure						
	Temporal Structures			Non-Temporal Structures			
	Future	Past	Perfect	Possessive	Evidentiality	Consonants	Colors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Crop Return (pre-1500CE)	0.12*** (0.03)	-0.06 (0.04)	0.05 (0.04)	-0.07* (0.04)	0.00 (0.03)	0.08 (0.06)	0.06 (0.34)
All Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted- R^2	0.11	0.08	0.14	0.15	0.20	0.31	-0.03
Observations	275	218	218	224	387	542	117

Notes: This table documents the positive, statistically, and economically significant association between pre-1500CE potential crop return and the existence of periphrastic future tense in a language, and not with other language structures. The analysis accounts for regional fixed-effects and other geographical characteristics as in previous tables. Other language structures include the existence a past tense, a perfect tense, the existence of obligatory possessive inflections, semantic distinctions of evidentiality, the number of consonants, the ratio of consonants to vowels and the number of colors. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S14.2: Crop Return and Tense Structure (OLS)

	Existence of Tense Structure			
	Periphrastic Future	Past Tense	Perfect Tense	Any Inflectional
	(1)	(2)	(3)	(4)
Crop Return (pre-1500CE)	0.15*** (0.04)	-0.07 (0.05)	0.06 (0.04)	0.06* (0.04)
Regional FE	Yes	Yes	Yes	Yes
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Semi-partial R^2 Crop Return (pre-1500CE)	0.06	0.01	0.01	0.02
Sum of Semi-partial R^2 Other Controls	0.14	0.16	0.19	0.14
Adjusted- R^2	0.09	0.10	0.12	0.06
Observations	173	173	173	173

Notes: This table documents the positive, statistically, and economically significant association between pre-1500CE potential crop return and the existence of periphrastic future tense in a language, and not with the presence of other tenses or the existence of inflectional tenses in a language. The analysis accounts for regional fixed-effects and other geographical characteristics as in previous tables. Heteroskedasticity robust standard error estimates clustered at the language family level are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S14.3: Geographical Origins of Gender and Language Structures

	Language Structure						
	Gender	Temporal Structures			Non-Temporal Structures		
		Past	Perfect	Possessive	Evidentiality	Consonants	Colors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Plow Negative Crops	-0.23***	-0.33***	-0.09	0.02	-0.05	-0.11	-0.76
(Average Caloric Yield, pre-1500)	(0.08)	(0.11)	(0.09)	(0.10)	(0.07)	(0.14)	(0.83)
All Crops	0.23***	0.24***	0.07	-0.06	0.02	0.08	0.73
(Average Caloric Yield, pre-1500)	(0.07)	(0.08)	(0.08)	(0.08)	(0.06)	(0.12)	(0.82)
All Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted- R^2	0.21	0.12	0.14	0.14	0.20	0.30	-0.03
Observations	217	218	218	224	387	542	117

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of sex-based grammatical gender in a language, and compares their impact on other language structures. The analysis accounts for regional fixed-effects and other geographical characteristics as in previous tables. Other language structures include the existence a past tense, a perfect tense, the existence of obligatory possessive inflections, semantic distinctions of evidentiality, the number of consonants, the ratio of consonants to vowels and the number of colors. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S14.4: Geographical Origins of Gender and Language Structures

	Language Structure						
	Gender	Temporal Structures			Non-Temporal Structures		
		Past	Perfect	Possessive	Evidentiality	Consonants	Colors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Plow Negative Crops	-0.22***	-0.33***	-0.09	0.01	-0.05	-0.09	-0.83
(Average Caloric Yield, pre-1500)	(0.08)	(0.11)	(0.09)	(0.10)	(0.07)	(0.14)	(0.83)
All Crops	0.25***	0.24***	0.07	-0.06	0.02	0.07	0.76
(Average Caloric Yield, pre-1500)	(0.07)	(0.08)	(0.08)	(0.08)	(0.06)	(0.12)	(0.82)
All Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted- R^2	0.24	0.12	0.14	0.14	0.20	0.31	-0.03
Observations	245	218	218	223	386	538	116

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of sex-based grammatical gender in a language, and compares their impact on other language structures. The analysis accounts for regional fixed-effects and other geographical characteristics as in previous tables. Other language structures include the existence a past tense, a perfect tense, the existence of obligatory possessive inflections, semantic distinctions of evidentiality, the number of consonants, the ratio of consonants to vowels and the number of colors. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S14.5: Geographical Origins of Gender and Language Structures

	Language Structure						
	Temporal Structures			Non-Temporal Structures			
	Gender	Past	Perfect	Possessive	Evidentiality	Consonants	Colors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Plow Negative Crops	-0.36**	-0.33***	-0.09	0.01	-0.05	-0.09	-0.83
(Average Caloric Yield, pre-1500)	(0.14)	(0.11)	(0.09)	(0.10)	(0.07)	(0.14)	(0.83)
Average Caloric Yield	0.39***	0.24***	0.07	-0.06	0.02	0.07	0.76
(All Crops, pre-1500)	(0.14)	(0.08)	(0.08)	(0.08)	(0.06)	(0.12)	(0.82)
All Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted- R^2	0.17	0.12	0.14	0.14	0.20	0.31	-0.03
Observations	181	218	218	223	386	538	116

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the intensity of grammatical gender in a language, and compares their impact on other language structures. The analysis accounts for regional fixed-effects and other geographical characteristics as in previous tables. Other language structures include the existence a past tense, a perfect tense, the existence of obligatory possessive inflections, semantic distinctions of evidentiality, the number of consonants, the ratio of consonants to vowels and the number of colors. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S14.6: Geographical Origins of Gender and Language Structures

	Language Structure						
	Gender	Temporal Structures			Non-Temporal Structures		
		Past	Perfect	Possessive	Evidentiality	Consonants	Colors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Plow Negative Crops	-0.18***	-0.33***	-0.09	0.01	-0.05	-0.09	-0.83
(Average Caloric Yield, pre-1500)	(0.07)	(0.11)	(0.09)	(0.10)	(0.07)	(0.14)	(0.83)
Average Caloric Yield	0.21***	0.24***	0.07	-0.06	0.02	0.07	0.76
(All Crops, pre-1500)	(0.06)	(0.08)	(0.08)	(0.08)	(0.06)	(0.12)	(0.82)
All Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted- R^2	0.13	0.12	0.14	0.14	0.20	0.31	-0.03
Observations	355	218	218	223	386	538	116

Notes: This table documents the statistically and economically significant association between the historical agricultural determinants of plow usage and the existence of gender distinctions in pronouns in a language, and compares their impact on other language structures. The analysis accounts for regional fixed-effects and other geographical characteristics as in previous tables. Other language structures include the existence a past tense, a perfect tense, the existence of obligatory possessive inflections, semantic distinctions of evidentiality, the number of consonants, the ratio of consonants to vowels and the number of colors. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S14.7: Geographical Origins of Politeness and Language Structures

	Language Structure						
	Politeness	Temporal Structures		Non-Temporal Structures			
		Past	Perfect	Possessive	Evidentiality	Consonants	Colors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ecological Diversity	0.10** (0.04)	0.04 (0.04)	0.05 (0.04)	-0.03 (0.04)	0.01 (0.03)	-0.10* (0.05)	-0.49 (0.37)
All Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted- R^2	0.32	0.07	0.14	0.14	0.20	0.31	-0.02
Observations	198	218	218	224	387	542	117

Notes: This table documents the positive, statistically, and economically significant association between ecological diversity and the existence of politeness distinctions in a language, and compares their impact on other language structures. The analysis accounts for regional fixed-effects and other geographical characteristics as in previous tables. Other language structures include the existence a past tense, a perfect tense, the existence of obligatory possessive inflections, semantic distinctions of evidentiality, the number of consonants, the ratio of consonants to vowels and the number of colors. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S14.8: Geographical Origins of Politeness and Language Structures

	Language Structure						
	Politeness	Temporal Structures		Non-Temporal Structures			
		Past	Perfect	Possessive	Evidentiality	Consonants	Colors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ecological Diversity	0.09*** (0.03)	0.04 (0.04)	0.05 (0.04)	-0.04 (0.04)	0.01 (0.03)	-0.08 (0.05)	-0.49 (0.38)
All Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted- R^2	0.25	0.07	0.14	0.15	0.20	0.31	-0.02
Observations	198	218	218	223	386	538	116

Notes: This table documents the positive, statistically, and economically significant association between ecological diversity and the existence of strong politeness distinctions in a language, and compares their impact on other language structures. The analysis accounts for regional fixed-effects and other geographical characteristics as in previous tables. Other language structures include the existence a past tense, a perfect tense, the existence of obligatory possessive inflections, semantic distinctions of evidentiality, the number of consonants, the ratio of consonants to vowels and the number of colors. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Table S14.9: Geographical Origins of Politeness and Language Structures

	Language Structure						
	Politeness	Temporal Structures		Non-Temporal Structures			
		Past	Perfect	Possessive	Evidentiality	Consonants	Colors
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ecological Diversity	0.23*** (0.07)	0.04 (0.04)	0.05 (0.04)	-0.04 (0.04)	0.01 (0.03)	-0.08 (0.05)	-0.49 (0.38)
All Geographic Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adjusted- R^2	0.32	0.07	0.14	0.15	0.20	0.31	-0.02
Observations	198	218	218	223	386	538	116

Notes: This table documents the positive, statistically, and economically significant association between ecological diversity and the existence of politeness distinctions in a language, and compares their impact on other language structures. The analysis accounts for regional fixed-effects and other geographical characteristics as in previous tables. Other language structures include the existence a past tense, a perfect tense, the existence of obligatory possessive inflections, semantic distinctions of evidentiality, the number of consonants, the ratio of consonants to vowels and the number of colors. Heteroskedasticity robust standard error estimates are reported in parentheses; *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level, all for two-sided hypothesis tests.

Figure S14.1: Geographical Origins and Other Language Structures (WALS)
Semi-partial R^2

(a) Pre-1500 Crop Return

(b) Plow Suitability

(c) Ecological Diversity

Figure S14.2: Geographical Origins and Other Language Structures (WALS)
Ratio of Semi-partial R^2 s

Figure S14.3: Orthogonality of Crop Return and Other Language Structures in WALs
Impact of Crop Return on Periphrastic Future Tense in Red

(a) Caloric Suitability of Plow Negative Crops and WALS Structures 1-68

(b) Caloric Suitability of Plow Negative Crops and WALS Structures 69-132

(c) Distribution of p -values

(d) Distribution of t -statistics

(e) Distribution of coefficients

Figure S14.4: Orthogonality of Suitability for the Usage of the Plow and Other Language Structures in WALS
(Impact of Determinants of Plow Usage on Sex-Based Grammatical Gender in Blue)

(a) Average Caloric Suitability and WALS Structures 1-68 (b) Average Caloric Suitability and WALS Structures 69-132

(c) Distribution of p -values (d) Distribution of t -statistics (e) Distribution of coefficients

Figure S14.4: Orthogonality of Suitability for the Usage of the Plow and Other Language Structures in WALS
(Impact of Determinants of Plow Usage on Sex-Based Grammatical Gender in Blue)

(f) WALS Structures 1-68

(g) WALS Structures 69-132

(h) Distribution of p -values

(i) Distribution of t -statistics

(j) Distribution of coefficients

Figure S14.5: Orthogonality of Ecological Diversity and Other Language Structures (WALS)
(Impact of Ecological Diversity on Politeness Distinctions in Orange)

Figure S14.6: Geographical Origins and Other Language Structures (WALS)
 Statistical Significance and Effect Sizes
 p -values and Absolute value of Estimated Coefficients (Logs vs Logs)

(a) Pre-1500 Crop Return

(b) Plow Suitability

(c) Ecological Diversity

Figure S14.7: Geographical Origins and Other Language Structures (WALS-Strict Binary Coding)

Statistical Significance and Effect Sizes

p -values and Absolute value of Estimated Coefficients (Logs vs Logs)

Figure S14.8: Geographical Origins and Other Language Structures (WALS-Binary Coding)
 Statistical Significance and Effect Sizes
 p -values and Absolute value of Estimated Coefficients (Logs vs Logs)

Figure S14.9: Geographical Origins and Other Language Structures (WALS-Binary Coding)
Common Sample
Statistical Significance and Effect Sizes
 p -values and Absolute value of Estimated Coefficients (Logs vs Logs)

Table S14.10: Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates

	Gender	Politeness	Future
	(1)	(2)	(3)
Coefficient	0.232	0.096	0.114
Uncorrected p-value	0.002	0.002	0.004
BH FDR corrected q-value	0.068	0.089	0.053
BKY sharpened FDR corrected q-value	0.072	0.095	0.052

Notes: This Table presents estimated coefficients and p -values based on the most complete specifications for each language structure in Table 1. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors clustered at the language family level. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.11: Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates
Binary Coding

	Gender	Politeness	Future
	(1)	(2)	(3)
Coefficient	0.232	0.096	0.114
Uncorrected p-value	0.002	0.002	0.004
BH FDR corrected q-value	0.064	0.070	0.042
BKY sharpened FDR corrected q-value	0.068	0.075	0.041

Notes: This Table presents estimated coefficients and p -values based on the most complete specifications for each language structure in Table 1. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors clustered at the language family level. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.12: Geographic Origins of Periphrastic Future Tense
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.114	0.114	0.114	0.114
Uncorrected p-value	0.001	0.001	0.001	0.001
BH FDR corrected q-value	0.008	0.006	0.011	0.010
BKY sharpened FDR corrected q-value	0.007	0.006	0.011	0.009
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	177	94	684	953

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALS that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, and (ii) there was any variation in the sample. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.13: Geographic Origins of Periphrastic Future Tense
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates (Clustered)

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.114	0.114	0.114	0.114
Uncorrected p-value	0.004	0.004	0.004	0.004
BH FDR corrected q-value	0.053	0.042	0.091	0.077
BKY sharpened FDR corrected q-value	0.052	0.041	0.096	0.080
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	177	94	684	953

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALs that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, and (ii) there was any variation in the sample. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors clustered at the family level. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.14: Geographic Origins of Periphrastic Future Tense
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates
Common Sample

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.182	0.182	0.182	0.182
Uncorrected p-value	0.000	0.000	0.000	0.000
BH FDR corrected q-value	0.004	0.004	0.009	0.007
BKY sharpened FDR corrected q-value	0.004	0.004	0.009	0.007
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	145	91	588	822

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALs that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, (ii) there was any variation in the sample, and (iii) sample was restricted to those languages that had information on the language structures of interest and one of the alternatives. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.15: Geographic Origins of Periphrastic Future Tense
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates
Common Sample (Clustered)

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.182	0.182	0.182	0.182
Uncorrected p-value	0.000	0.000	0.000	0.000
BH FDR corrected q-value	0.001	0.001	0.002	0.002
BKY sharpened FDR corrected q-value	0.001	0.001	0.002	0.002
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	145	91	588	822

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALs that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, (ii) there was any variation in the sample, and (iii) sample was restricted to those languages that had information on the language structures of interest and one of the alternatives. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors clustered at the family level. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.16: Geographic Origins of Sex-Based Grammatical Gender
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.232	0.232	0.232	0.232
Uncorrected p-value	0.004	0.004	0.004	0.004
BH FDR corrected q-value	0.112	0.098	0.122	0.127
BKY sharpened FDR corrected q-value	0.126	0.106	0.137	0.145
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	177	94	684	953

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALs that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, and (ii) there was any variation in the sample. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.17: Geographic Origins of Sex-Based Grammatical Gender
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates (Clustered)

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.232	0.232	0.232	0.232
Uncorrected p-value	0.002	0.002	0.002	0.002
BH FDR corrected q-value	0.068	0.064	0.120	0.116
BKY sharpened FDR corrected q-value	0.072	0.068	0.137	0.131
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	177	94	684	953

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALS that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, and (ii) there was any variation in the sample. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors clustered at the family level. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.18: Geographic Origins of Sex-Based Grammatical Gender
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates
Common Sample

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.288	0.288	0.288	0.288
Uncorrected p-value	0.029	0.029	0.029	0.029
BH FDR corrected q-value	0.346	0.261	0.401	0.379
BKY sharpened FDR corrected q-value	0.474	0.316	0.643	0.566
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	145	91	588	822

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALS that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, (ii) there was any variation in the sample, and (iii) sample was restricted to those languages that had information on the language structures of interest and one of the alternatives. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.19: Geographic Origins of Sex-Based Grammatical Gender
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates
Common Sample (Clustered)

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.288	0.288	0.288	0.288
Uncorrected p-value	0.042	0.042	0.042	0.042
BH FDR corrected q-value	0.395	0.347	0.440	0.439
BKY sharpened FDR corrected q-value	0.580	0.464	0.699	0.693
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	145	91	588	822

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALs that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, (ii) there was any variation in the sample, and (iii) sample was restricted to those languages that had information on the language structures of interest and one of the alternatives. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors clustered at the family level. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.20: Geographic Origins of Politeness Distinctions
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.096	0.096	0.096	0.096
Uncorrected p-value	0.011	0.011	0.011	0.011
BH FDR corrected q-value	0.125	0.212	0.232	0.206
BKY sharpened FDR corrected q-value	0.132	0.264	0.285	0.245
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	177	94	684	953

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALs that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, and (ii) there was any variation in the sample. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.21: Geographic Origins of Politeness Distinctions
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates (Clustered)

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.096	0.096	0.096	0.096
Uncorrected p-value	0.002	0.002	0.002	0.002
BH FDR corrected q-value	0.089	0.070	0.148	0.142
BKY sharpened FDR corrected q-value	0.095	0.075	0.172	0.163
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	177	94	684	953

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALS that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, and (ii) there was any variation in the sample. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors clustered at the family level. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.22: Geographic Origins of Politeness Distinctions
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates
Common Sample

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.154	0.154	0.154	0.154
Uncorrected p-value	0.004	0.004	0.004	0.004
BH FDR corrected q-value	0.128	0.070	0.256	0.202
BKY sharpened FDR corrected q-value	0.143	0.072	0.341	0.249
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	145	91	588	822

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALS that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, (ii) there was any variation in the sample, and (iii) sample was restricted to those languages that had information on the language structures of interest and one of the alternatives. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Table S14.23: Geographic Origins of Politeness Distinctions
Robustness to Multiple Hypothesis Testing
Accounting for False Discovery Rates
Common Sample (Clustered)

	All	Binary	All Binary	Full
	(1)	(2)	(3)	(4)
Coefficient	0.154	0.154	0.154	0.154
Uncorrected p-value	0.004	0.004	0.004	0.004
BH FDR corrected q-value	0.104	0.078	0.278	0.195
BKY sharpened FDR corrected q-value	0.115	0.082	0.385	0.243
Main Geographical Controls	Yes	Yes	Yes	Yes
Precipitation Controls	Yes	Yes	Yes	Yes
Temperature Controls	Yes	Yes	Yes	Yes
Unproductive Period	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes
Number of Hypotheses Tested	145	91	588	822

Notes: This Table presents estimated coefficients and p -values based on the most complete specification in Table 1. FDR correction based on running alternative regressions on all variables of the WALs that satisfied the following conditions: (i) the sample in the regression included at least 50 languages, (ii) there was any variation in the sample, and (iii) sample was restricted to those languages that had information on the language structures of interest and one of the alternatives. Uncorrected and corrected p -values are based on heteroskedasticity robust standard errors clustered at the family level. False Discovery Rates (FDR) corrected q -values based on the methods presented in (149) (BH) and (150) using the implementation of (151). q -values have similar interpretation to p -values and give the probability threshold at which the Null hypothesis would be rejected after accounting for multiple hypothesis testing (see (151)).

Appendix References

49. M. P. Lewis, G. F. Simons, C. D. Fennig, *Ethnologue: Languages of the world*, vol. 16 (SIL international Dallas, TX, 2009).
50. G. P. Murdock, *Ethnology* pp. 109–236 (1967).
51. G. P. Murdock, D. R. White, *Ethnology* pp. 329–369 (1969).
52. S. Wichmann, A. Müller, V. Velupillai, *Diachronica* **27**, 247 (2010).
53. M. Dryer, Matthew S & Haspelmath, ed., *The World Atlas of Language Structures Online*. (Leipzig: Max Planck Institute for Evolutionary Anthropology., 2013).
54. Ö. Dahl, *Tense and aspect systems* (Basil Blackwell, 1985).
55. Ö. Dahl, *Tense and Aspect in the Languages of Europe* p. 309 (2000).
56. Ö. Dahl, V. Velupillai, *The World Atlas of Language Structures Online*, M. S. Dryer, M. Haspelmath, eds. (Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013).
57. J. L. Bybee, W. Pagliuca, *Papers from the 7th international conference on historical linguistics* (Amsterdam: John Benjamins, 1987), pp. 109–122.
58. J. L. Bybee, Ö. Dahl, *Studies in language* **13**, 51 (1989).
59. J. L. Bybee, R. Perkins, W. Pagliuca, *The evolution of grammar* (1994).
60. D. Stahlberg, F. Braun, L. Irmen, S. Sczesny, *Social communication* pp. 163–187 (2007).
61. R. Lakoff, *Language in society* **2**, 45 (1973).
62. P. Eckert, S. McConnell-Ginet, *Language and gender* (Cambridge University Press, Cambridge, 2003).
63. G. G. Corbett, *The World Atlas of Language Structures Online*, M. S. Dryer, M. Haspelmath, eds. (Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013).
64. A. Siewierska, *Gender Distinctions in Independent Personal Pronouns* (Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013).
65. G. G. Corbett, *Number of Genders* (Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013).
66. G. G. Corbett, *Systems of Gender Assignment* (Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013).

67. R. Brown, A. Gilman, *Style in Language*, T. A. Sebeok, ed. (MIT Press, 1960), chap. 12.
68. Y. R. Solé, *Thesaurus* **1**, 161 (1970).
69. N. del Castillo Mathieu, *Thesaurus. Boletín del Instituto Caro y Cuervo Bogota* **37**, 602 (1982).
70. P. Bentivoglio, *Diachronic Perspectives on Address Term Systems*, I. Taavitsainen, A. H. Jucker, eds. (John Benjamins Publishing, 2003), pp. 177–192.
71. J. Helmbrecht, *The World Atlas of Language Structures Online*, M. S. Dryer, M. Haspelmath, eds. (Max Planck Institute for Evolutionary Anthropology, Leipzig, 2013).
72. O. Galor, Ö. Özak, *American Economic Review* **106**, 3064 (2016).
73. F. L. Pryor, *Comparative Studies in Society and history* **27**, 727 (1985).
74. A. Alesina, P. Giuliano, N. Nunn, *The Quarterly Journal of Economics* **128**, 469 (2013).
75. J. Fenske, *Journal of the European Economic Association* **12**, 612 (2014).
76. D. M. Olson, *et al.*, *BioScience* **51**, 933 (2001).
77. G. H. Hofstede, *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (Sage, 2001).
78. M. G. Marshall, T. R. Gurr, K. Jaggers, *Center for Systemic Peace* (2014).
79. A. Chanda, L. Putterman, *States and Development* (Springer, 2005), pp. 69–91.
80. A. Falk, *et al.*, *The Quarterly Journal of Economics* **133**, 1645 (2018).
81. A. Falk, A. Becker, T. Dohmen, D. Huffman, U. Sunde (2016).
82. R. J. Barro, J. W. Lee, *Journal of development economics* **104**, 184 (2013).
83. C. Welzel, *Freedom rising: human empowerment and the quest for emancipation* (Cambridge University Press, New York, 2013).
84. S. Wichmann, E. W. Holman, *Temporal stability of linguistic typological features* (Lincom Europa, 2009).
85. S. J. Greenhill, Q. D. Atkinson, A. Meade, R. D. Gray, *Proceedings of the Royal Society of London B: Biological Sciences* **277**, 2443 (2010).
86. D. Dediu, S. C. Levinson, *PloS One* **7**, e45198 (2012).
87. D. Dediu, M. Cysouw, *PloS one* **8**, e55009 (2013).

88. S. J. Greenhill, *et al.*, *Proceedings of the National Academy of Sciences* **114**, E8822 (2017).
89. D. Figlio, P. Giuliano, U. Özek, P. Sapienza, Long-term orientation and educational performance, *Tech. rep.*, National Bureau of Economic Research (2016).
90. A. Falk, *et al.*, *The Quarterly Journal of Economics* **133**, 1645 (2018).
91. J. G. Altonji, T. E. Elder, C. R. Taber, *Journal of Political Economy* **113**, 151 (2005).
92. E. Oster, Unobservable selection and coefficient stability: Theory and validation (2014).
93. S. Lestrade, H. de Hoop, *The Linguistic Review* **33**, 397 (2016).
94. H. Reichenbach, *Time: From Concept to Narrative Construct: a Reader* (1947).
95. P. S. Baker, *Introduction to Old English* (John Wiley & Sons, 2012).
96. E. Boserup, *Woman's role in economic development* (St. Martin's Press, New York, 1970).
97. E. Depetris-Chauvin, Ö. Özak, *Working Paper, Southern Methodist University, Department of Economics* (2016).
98. P. Brown, S. C. Levinson, *Politeness: Some universals in language usage*, vol. 4 (Cambridge university press, 1987).
99. M. Sifianou, *Politeness phenomena in England and Greece: a cross-cultural perspective* (Clarendon Press, Oxford, 1992).
100. R. L. Carneiro, *Science* **169**, 733 (1970).
101. T. G. Conley, *Journal of econometrics* **92**, 1 (1999).
102. GLOBE Task Team and others, ed., *The Global Land One-kilometer Base Elevation (GLOBE) Digital Elevation Model, Version 1.0*. (National Oceanic and Atmospheric Administration, National Geophysical Data Center, 325 Broadway, Boulder, Colorado 80303, U.S.A., 1999).
103. S. Riley, S. DeGloria, R. Elliot, *Intermountain Journal of Sciences* **5**, 23 (1999).
104. R. Durante (2010).
105. S. Wichmann, *et al.*, *The ASJP Database (Version 12)* (2009).
106. S. Wichmann, E. W. Holman, C. H. Brown, eds., *The ASJP Database (version 18)* (2018).
107. C. Ehret, *Examining the farming/language dispersal hypothesis* pp. 163–176 (2002).

108. D. Denoon, G. McCormack, *Multicultural Japan: palaeolithic to postmodern* (Cambridge University Press, 2001).
109. M. Mithun, *The languages of native North America* (Cambridge University Press, 2001).
110. P. Aalto, *Central Asiatic Journal* pp. 9–16 (1955).
111. L. Campbell, M. Mithun, *The languages of native America: Historical and comparative assessment* (University of Texas Press, 2014).
112. W. F. Adelaar, *The languages of the Andes* (Cambridge University Press, 2004).
113. J. Diamond, P. Bellwood, *Science* **300**, 597 (2003).
114. P. SUTTON, H. KOCH, R. Dixon, *Journal of Linguistics* **44**, 471 (2008).
115. P. Sidwell, R. Blench, *Dynamics of human diversity* p. 315 (2011).
116. P. Bellwood, *Prehistory of the Indo-Malaysian Archipelago: revised edition* (ANU Press, 2013).
117. A. Chowning, *Pacific Linguistics. Series A. Occasional Papers* p. 17 (1969).
118. C. Voorhoeve .
119. L. Campbell, *American Indian languages: the historical linguistics of Native America* (Oxford University Press, 2000).
120. M. Donohue, *et al.*, *Current Anthropology* **51**, 223 (2010).
121. B. Krishnamurti, *The dravidian languages* (Cambridge University Press, 2003).
122. A.-A. Abu-Manga, L. G. Gilley, A. Storch, *Insights Into Nilo-Saharan Language, History and Culture: Proceedings of the 9th Nilo-Saharan Linguistics Colloquium, Institute of African and Asian Studies, University of Khartoum, 16-19 February 2004*, vol. 23 (Köppe, 2006).
123. W. A. Foley, *Annual review of anthropology* **29**, 357 (2000).
124. T. G. Schurr, *et al.*, *American journal of physical anthropology* **148**, 422 (2012).
125. R. Blench, M. Dendo, *Human migrations in continental East Asia and Taiwan. Matching archaeology, linguistics and genetics* pp. 105–132 (2008).
126. D. W. Anthony, *The horse, the wheel, and language: how Bronze-Age riders from the Eurasian steppes shaped the modern world* (Princeton University Press, 2010).

127. R. A. Miller, *Languages and History: Japanese, Korean, and Altaic* (White Orchid Press, 1996).
128. C. H. Brown, C. R. Clement, P. Epps, E. Luedeling, S. Wichmann, *Ethnobiology Letters* **4**, 1 (2013).
129. M. Clendon, *et al.*, *Current Anthropology* **47**, 39 (2006).
130. P. Manning, *Journal of World History* **17**, 115 (2006).
131. J. M. Diamond, *Guns, germs, and steel: the fates of human societies* (W.W. Norton & Co., New York, 1997), first edn.
132. M. C. Campbell, S. A. Tishkoff, *Current biology* **20**, R166 (2010).
133. V. Chaussonnet, *Crossroads Alaska: native cultures of Alaska and Siberia* (Arctic studies center, National museum of natural history, Smithsonian institution, 1995).
134. E. W. Campbell, *Language and Linguistics Compass* **11**, e12240 (2017).
135. J. G. Pena, A historical reconstruction of the peba-yaguan linguistic family, Ph.D. thesis, University of Oregon (2009).
136. R. J. LaPolla, *Linguistics of the Tibeto-Burman Area* **39**, 282 (2016).
137. P. Bellwood, *Scientific American* **265**, 88 (1991).
138. R. Blench, L. Sagart, A. Sanchez-Mazas, *The Peopling of East Asia* (Routledge, 2005), pp. 135–159.
139. L. Eriksen, A. V. Galucio, *The native languages of South America: origins, development, typology* pp. 177–202 (2014).
140. J. Janhunen, *The quasiquicentennial of the Finno-Ugrian society* **258**, 57 (2009).
141. M. Durbin, *South American Indian Languages. Retrospect and Prospect* **325**, 370 (1985).
142. P. Trudgill, *Linguistic Typology* **8**, 305 (2004).
143. C. Uhlenbeck, *Jukagirisch und uralisch* (1940).
144. R. Lal, D. Reicosky, J. Hanson, *Soil & Tillage Research* **1**, 1 (2007).
145. R. Lal, *Journal of sustainable agriculture* **33**, 66 (2009).
146. R. Blench, *Archaeology of African Plant Use* **61**, 243 (2013).

147. S. Piggott, *The earliest wheeled transport, from the Atlantic coast to the Caspian Sea* (1983).
148. J. Pokorny, *Indogermanisches etymologisches Wörterbuch* (Francke, Bern, 1948).
149. Y. Benjamini, Y. Hochberg, *Journal of the Royal statistical society: series B (Methodological)* **57**, 289 (1995).
150. Y. Benjamini, A. M. Krieger, D. Yekutieli, *Biometrika* **93**, 491 (2006).
151. M. L. Anderson, *Journal of the American statistical Association* **103**, 1481 (2008).