Do Languages Generate Future-Oriented Economic Behavior?

Experimental Evidence for Causal Effects

Abstract

Studies have shown that the use of languages which grammatically associate the future and the present tends to correlate with more future-oriented behavior. We take an experimental approach to go beyond correlation. We asked bilingual research participants, people fluent in two languages (12 language pairs) which differ in the way they encode time, to make a set of future-oriented economic decisions. We find that participants addressed in a language in which the present and the future are marked more distinctly tended to value future events less than participants addressed in a language in which the present and the future are similarly marked. We supplement the analysis with a within-person experiment in which bilingual research participants (8 language pairs) were asked to spatially mark the distance between the present and the future. When participants were addressed in a language in which the present and the future are marked more distinctly they tended to view the distance between the present and the future as greater, compared to when addressed in a language in which the present and the future are marked less distinctly.

Ian Ayres, Yale Law School ian.ayres@yale.edu

Tamar Kricheli-Katz, Tel Aviv University, Law School, tamarkk@post.tau.ac.il

Tali Regev, Reichman University (IDC), Economics Department, tregev@idc.ac.il

Do languages affect or merely reflect the attitudes, preferences, and behaviors of the people who speak them?

Numerous studies have documented correlations between the linguistic features and grammatical structures of languages and the attitudes, preferences, and behaviors of the people who speak them. Thus, for example, it has been shown that speakers of languages with different structures and features differ accordingly in their processing of colors, future-oriented economic behaviors, and gendered attitudes (Chen 2013, Davis and Reynolds 2018, Gay et al. 2013, Jakiela and Ozier 2018, Mavisakalyan et al. 2018, Prewitt-Freilino et al. 2012, Galor et al. 2020, Ladd et al. 2015). However, evidence for the causal effects of the features and structure of languages on attitudes, preferences, and behaviors is harder to establish. It is a challenge to demonstrate empirically that using a specific language can affect, and not just merely reflect, the way we perceive the world. Indeed, scholars across several disciplines have debated, and continue to debate, the relationship between language and thought (Ladd et al. 2015). Whereas some have argued that languages do not restrict people's perceptions and behavior (Li and Gleitman 2002), others (who subscribe to the linguistic relativity hypothesis) have asserted that speakers of languages develop language-specific schemas and structures which affect their perceptions and behavior (Whorf 1956, Carroll et a. 2004, Levinson 2012, Lucy 1996).

We contribute to this longstanding debate by providing evidence for the causal impact of the encoding of time in the language spoken on the intertemporal economic choices that people make and on their perceptions of the distance between the present and the future. Our findings suggest that perceptions of time are differently embedded in languages, and can impact everyday human behavior.

Languages vary in the ways in which they encode time (Dahl 2000,Thieroff 2000, Chen 2013). In some languages, like German, the same grammatical tense can be used to refer to both present and future events ("futureless languages"; "weakly-grammaticalized future" languages "; "weak FTR languages"). Other languages, like English and French have the obligatory grammatical marking of the future tense. (“futured languages”; “strong FTR languages”).[1]

Studies based on survey data show that the use of futureless languages, which grammatically associate the future and the present, tends to correlate with more future-oriented behavior on the part of people and organizations. Thus, for example, across and within countries, speakers of such languages save more, retire with more accumulated wealth, smoke less, practice safer sex, are less obese, and care more about the environment (Chen 2013, Mavisakalyan et al. 2018, Liang et al. 2018, Hübner and Vannoorenberghe 2015, Fasan et al. 2016, Chen et al. 2015, Campo et al. 2020, Kim et al. 2017, Karapandza 2016).

Why would speakers of weak FTR languages express more future-oriented economic behavior?

One possible explanation is that consistently speaking about future events in the present tense can make the future seem more immediate. Thus, the speakers of such languages may tend to value future rewards more than the speakers of languages in which the present and the future are distinctly marked (Chen 2013). Another possible explanation is that because weak-FTR languages do not obligate speakers to mark present and future events differently, speakers of these languages might not think as precisely about the temporal distance of future events as speakers of strong-FTR languages. As a result, speakers of weak-FTR languages may tend to discount future rewards less than speakers of strong-FTR languages (Chen 2013).

Another possible explanation for the correlation found between languages and future-oriented economic behaviors (Chen 2013) is that cultural differences regarding time preferences across and within countries (Costa-Font et al. 2018) might be reflected both

[1] It should be noted that almost all strong-FTR languages have some uses in which marking the future is not necessary; but these are mostly non-canonical or minor uses (Dahl 2000). Thus, for example, in English (a strong FTR) one can say “I’m leaving tomorrow” or “the term begins next week.”

in the languages spoken and in the observed differences in the speakers' future economic behaviors.

Although survey data does provide the opportunity to identify correlations between the way in which a language encodes time and the future-oriented economic behaviors of the people who speak it, it is hard to use them as a means of determining a *causal effect* between the language spoken and future-oriented economic behaviors. In other words, it is nearly impossible to hold constant the unmeasured cultural differences, across and within countries, that might be reflected both in the language spoken and in the observed differences in the future economic behaviors of its speakers. Indeed, following Chen's (2013) research, some studies have argued that the research merely shows that the languages we speak reflect the societies and cultures in which we live, but does not show that the languages we speak influence our perceptions and behaviors (Pullum 2012, Grossman 2013, Liberman 2013, see also: Roberts et al. 2015).

Weighing in on the debate, Sutter and colleagues (2018) studied differences in the intertemporal choices of children living in a bilingual city where about half of the inhabitants spoke German (a weak-FTR language) and the other half spoke Italian (a strong-FTR language). Using an intertemporal choice experiment, they found that German-speaking primary school children were more likely than their Italian-speaking peers to delay gratification. However, like Chen's study, their study could not rule out the possibility that the observed differences in behavior were generated by the cultural differences between the two groups. Galor et al. (2018) sought to provide evidence for the causal effect of the encoding of time in a language on its speakers' educational attainment. To do so, they exploited variations in the native languages spoken by the children of migrants with identical ancestral countries of origin; they were able to show the significant positive effect of speaking a futureless language on educational attainment. Because the analysis compared children with the same ancestral countries of origin, the experiment's design controlled for all the cultural differences associated with countries of origins. Yet, it is not possible to fully control for the cultural differences associated with sub-cultures within countries (those reflected in languages within countries).

We take a different methodological approach to identifying the *causal effect* of language on future-oriented behavior. We wish to show that the encoding of time in a language not only reflects, but also generates differences in future-oriented economic behaviors. Thus, we hypothesize that when people are addressed in a futured language, one that grammatically differentiates between the present and the future—i.e., it has a strong FTR—they tend discount future events more. This, in turn, encourages less future-oriented behavior, such as spending more in the present. When people are addressed in a futureless language, one that does not differentiate grammatically between the present and the future—i.e., it has a weak FTR—they discount the future less. This, in turn, encourages more future-oriented behavior, such as saving for the future.

We also wish to show that perceptions of time are differently embedded in languages, and are activated when the language is spoken. Thus, we hypothesize that that when people are addressed in a futured language (a strong FTR language) they tend to view the future as more distant compared to when addressed in a futureless language (a weak FTR language). To do so, we used two experimental designs: a randomized between-subject experimental design and a randomized within-subject experimental design. In the first study, we asked bilingual people, fluent in two languages which differ in the way that they encode time, to make a future-oriented economic decision: specifically, we asked participants, in one of the two languages in which they are fluent, to make a set of binary choices about whether they wished to be paid a certain amount of money earlier (today), or a larger amount of money later (the following week). We then tested whether the people randomly assigned the question in a strong-FTR language required more by way of future compensation than those asked the question in a weak-FTR language.

Following the EUROTYP project (Dahl 1985, Dahl 2000) and Chen (2013), we separated the languages we explored into two broad categories: weak- and strong-FTR. Strong-FTR languages are those that require future events to be grammatically marked when making predictions. Weak-FTR languages do not require such grammatical marking.

Altogether, in the first study we used 12 language pairs, in which one language uses the same grammatical tense for the present and the future—i.e., it has a weak future time reference (weak FTR)—and the other has a strong future time reference (strong FTR).[2]

In the second study, we directly asked bilingual people, fluent in two languages which differ in the way in which they encode time, to mark the distance between the present and the future: specifically, we asked participants, in one of the two languages in which they are fluent to spatially mark the distance between 'now' and 'later' and then in the other language to spatially mark the distance between 'today' and 'tomorrow'. Altogether, in the second study we used eight language pairs, in which one language has a weak future time reference (weak FTR)—and the other has a strong future time reference (strong FTR).

The Experiments

Study 1

The experiment involved bilingual participants, proficient in one weak-FTR language (German, Dutch, Mandarin) and one strong-FTR language (English, French, Spanish, Hindi).

The experiment was conducted in the spring and summer of 2019. Participants, recruited via MTurk (a crowdsourcing marketplace for Human Intelligence Tasks (HIT)), were randomly assigned to either the weak- or strong-FTR experimental condition. They were first asked (in either the randomly assigned weak- or strong-FTR language) to make a set of binary choices about whether they wished to be paid a certain amount of money earlier ($3 today), or a larger amount of money later (the following week). We asked the participants to make a choice with eleven binary decision problems constructed thus, varying the value of the future compensation to be received a week hence from $3.05 to $7. This procedure builds upon previous studies, in which similar multiple price

hypotheses were described before running the study in an IRB application that is available upon request.

list procedures were used to elicit the participants' discount rates (i.e., the amount for which participants were willing to receive a delayed payment (Daniel et al. 2010, Frederick et al. 2002). After choosing their preferred payment methods, participants were asked about their level of comfort in both languages, and their country of residence. They were then asked to take language proficiency tests in the two languages in which they had declared themselves to be proficient. Each language proficiency test consisted of nine questions. The order of the two proficiency tests was randomized. (See the SI Appendix for the payment options and proficiency tests). After completing the two proficiency tests, participants were then asked a series of demographic questions. The geolocation of the participants was also coded. Participants were then given instructions on how to receive payment. The results of participants who proved not to be proficient in both tested languages were excluded from the study.

The experiment therefore consisted of 12 sub-experiments (4 Strong-FTR languages x 3 Weak-FTR languages) X 2 experimental conditions (Compensation Question in Strong-FTR or Weak-FTR language). Over the course of eight months starting in November 2018, we published 12 different HITs on MTurk, inviting bilingual participants to participate in our study if they were genuinely fluent in the two languages.

Altogether 6,189 participants declared themselves to be bilingual and fluent in the two languages, but only 3,804 completed the experiment. Only 717 of the participants passed the two language proficiency tests assigned to them (i.e., received a score of at least 6 out of 9 in each of the proficiency tests). We also excluded participants who displayed inconsistent time preferences, and those who participated from the same IP address as other participants. The final sample used in the analysis consists of 565 participants (see Table 1).

(Table 1 about here)

Of the 565 participants, 289 were assigned to the weak-FTR condition and 276 to the strong-FTR condition. Table 2 presents the sample characteristics, by the experimental condition.

(Table 2 about here)

Results

Table 3 presents the results of Tobit regression models predicting participants' lowest accepted delayed payment value. For each participant, we capture the lowest amount for which she indicated a preference to be paid a week from now, rather than being paid $3 immediately. Participants who provided inconsistent time preferences were excluded from the analysis.

We use Tobit (censored) regression models because the dependent variable is left and right censored; our sample includes participants for whom it was impossible to determine their precise preferences (those who denied all delayed payment offers (31.2% of participants), and those who accepted all delayed payment offers (7.4%)).

(Table 3 about here)

In Model 1, we estimate the effect of the experimental condition, controlling for the participants' proficiency in the language in which they were asked the payment questions. In Model 2, we add the gap in the participants' proficiency in the strong, compared to the weak, FTR language. The gap reflects the participants' relative immersion in the strong, compared to the weak, FTR language and culture, and captures the correlations observed by Chen (2013). Model 3 includes an additional interaction between the gap in proficiency and the experimental conditions. Model 4 includes the demographic characteristics of the participants, and the language pair in which the participants were bilingual. Finally, following Roberts et al. (2015), Model 5 controls for the origins of the languages in which the participants are fluent (and thus does not include the specific language pair).

As predicted, being addressed in the strong-FTR language generated a higher time discount rate than being addressed in the weak-FTR language.

In Models 1,2,4, and 5, the lowest accepted delayed payment for participants addressed in a strong FTR language was at least 50 cents higher than the lowest accepted delayed payment for participants who were addressed in the weak FTR ($p<0.001$). The effects of being addressed in the strong-FTR language, in all the models, are statistically significant. Participants' proficiency in the language in which the payment questions were asked also affected their preferences. More proficient participants had lower time discount rates, compared to the less proficient participants. This may be because the more proficient participants understood the payment questions better, or because more fluent participants are also more willing to delay immediate reward, compared to less proficient participants (scores in the two proficiency tests were positively correlated).

The strong-weak gap variable in Models 2 and 3 captures participants' relative proficiency in the strong, compared to the weak, FTR language. It thus captures the relative immersion of participants in the strong FTR language and culture. The significant and positive effect in Model 2 ($p<0.05$) implies that participants who are relatively more proficient in the strong-FTR language (compared to their proficiency in the weak-FTR language) tend to have a higher time discount rate, compared to participants who are relatively more fluent in the weak-FTR language.

The significant and positive interaction in Model 3 (asked in the strong-FTR * strong-weak gap, $p<0.001$) suggests that the effects of being asked the payment questions in the strong-FTR language are significantly stronger for participants who are more proficient in the strong-FTR language.

Finally, for robustness, we replicated the analyses while controlling for the participants' self-reported native language (in additions to the proficiency gap we observed). We also replicated the analyses with country of residence fixed effects (so as to control for the differences in the impact of $1 on the participants' wellbeing). Results remained similar in magnitude and statistical significance.

In the SI Appendix (Tables S1 and S2), we present the results of OLS and ordered-logit regression models predicting participants' reservation prices. The results of these models are similar to the results presented in Table 3.

Attrition and Selection Bias

Only 61% of the participants who started the experiment completed it (see Table 1). Our concern is that participants who were asked the payment questions in a language in which they were less proficient tended to quit the study more; which is to say, participants who were more fluent in the strong-FTR language and were asked the payment questions in the weak-FTR language were more likely to leave the experiment than those who were more fluent in the strong-FTR language and were asked the payment questions in the strong-FTR language.

If this indeed was the case, the sample would generate a biased sample, in which the participants who were asked the questions in the strong FTR-language would also be more proficient in the strong-FTR language, compared to the participants who were asked the questions in the weak-FTR language. If the sample were indeed biased, it would be impossible to disentangle the effects of the experimental condition (being asked in the strong-FTR language) from the effect of one's proficiency in the strong-FTR language.

In Table S3, we present the results of balancing tests, comparing the characteristics of participants by experimental condition.

We can see that participants who were asked the payment questions in the strong-FTR language are not significantly more fluent in the strong-FTR language, compared to the participants who were addressed in the weak-FTR language. Yet, the findings suggest that participants who were fluent in French (a strong-FTR language) were disproportionately represented in the strong-FTR experimental condition. To eliminate the concern that this imbalance generated the results that we observed, we estimated the same Tobit regression models predicting participants' lowest accepted delayed payment after excluding all the French-speaking participants from the sample. The results we obtained are very similar to the results obtained with the full sample (see Table S4 in the SI Appendix). This suggests that it was not the attrition from the experiment that generated the results we report. Finally, the results of the balancing tests presented in Table 5 further suggest that people who were fluent in Hindi (a strong-FTR language) were disproportionately *underrepresented* in the strong-FTR condition. Although this bias should decrease the probability of obtaining the results we report, we also estimated the same Tobit regression models predicting participants' lowest accepted delayed payment after excluding all the Hindi-speaking participants from the sample. The results we obtained are not statistically different to the results presented in Table 2 (see Table S2 in the SI Appendix).

The English Language

Following Chen (2013), we coded the English language as having a strong FTR. However, critics have argued that English does not have an obligatory grammatical marking of future events (Pullum 2012). To eliminate the concern that participants who were addressed in English generated the results we observed, we estimated the same Tobit regression models predicting participants' lowest accepted delayed payment after excluding all the English-speaking participants from the sample. The results we obtained are very similar to the results obtained with the full sample (see Table S5 in the Appendix).

Following this concern, and to rule out the possibility that any other language or pair of languages may have driven the results we observed, we repeated the analyses on sub-samples of the data, excluding one language or pair of languages at a time. The effects remained similar in magnitude and statistical significance.[3] We therefore conclude that the effects we observed are not generated by one of the languages or by one pair of the languages that we studied.

Our first study has some limitations. Most notably, it is possible that the languages we chose are associated with cultural scripts regarding trust, the strength of the economy, or norms about savings (Falk et al. 2018). Thus for example, the strong-FTR languages we chose (English, French, Spanish, Hindi) might be associated with countries with less wealthy economies than the countries associated with the weak-FTR languages we chose (Mandarin, Dutch, German). If this is the case, being addressed in a strong-FTR language would activate beliefs about weaker economies and lower probabilities of being paid in the future and thus lead participants to trust future payments less.

[3] These results can be found in our Online Appendix.

To eliminate this concern we reran our analysis on a sample which excluded Hindi and Spanish speakers. Results remained similar in magnitude and statistical power.

A related concern is that the encoding of time in a language is correlated with the encoding of notions of probability (Nuyts, 2000; Palmer, 2001). As a result, being addressed in a strong-FTR language activates not only scripts about the distance between the present and the future but also scripts about the probability of being paid in the future. In addition, in order to reduce this and related concerns and to provide direct evidence for the mechanism that generated the results we observed, we supplemented the analysis with an additional within person experiment. The second experiment directly explores the effects of the encoding of time in a language on the perceived distance between the present and the future.

Study 2

The second experiment tested whether people viewed the future as being more distant when addressed in a strong FTR language compared to a weak FTR language. It involved bilingual participants, proficient in one weak-FTR language (German, Dutch, Mandarin) and one strong-FTR language (English, French, Spanish[4]). The experiment was conducted in the spring of 2022. Participants were targeted via Facebook ads and directed to our Qualtrics survey. The experiment involved two stages, one in each language (the weak- or strong-FTR language). Each stage involved first taking a proficiency test and then reporting the perceived distance between the present and the future. The language proficiency tests in the experiment were used both to test for participants' proficiency and to make the language salient. All the participants were asked

[4]We also tried to recruit participants who were fluent in Hindi (and in additional weak FTR language), but were not able to recruit enough bilingual research participants who were fluent in Hindi as well as in Mandarin, Dutch or German.

questions in the two languages in which they declared being proficient (A within subject design).

After being directed to our survey, participants were randomly assigned to whether questions were first asked in the weak- or strong-FTR language. Participants were first asked to take a language proficiency test in one language (either the weak or strong FTR) and were then asked (in the first language), using a slider, to spatially mark the distance between the present and the future (participants were randomly assigned to mark the distance between 'now' and 'later' or between 'today' and 'tomorrow'). Figure 1 demonstrates the sliders used in our experiment and the average responses of participants (in English).

As presented in the figure, the sliders we used did not include visible measurement units. Yet, we capture the perceived distance between the present and the future on a continuous variable (0-1) which represents the distance between the present and the future.

After completing the first stage participants were asked to take a second language proficiency test in the second language in which they declared being fluent (either the weak or strong FTR-language based on the random assignment). They were then asked, again, using a slider, to spatially mark the distance between the present and the future. Participants who were randomly assigned in the first stage to mark the distance between 'now' and 'later' were asked in the second stage to mark the distance between 'today' and 'tomorrow', and vice versa.[5]

The final sample for the study included 570 participants who received scores higher than 6/9 in the two proficiency tests they took (participants with lower proficiency scores were removed from the sample). In table S6 in the appendix, we report the sample characteristics. In table S7 we report the number of participants in each language-pair we

[5] Hypotheses and design were registered at https://osf.io/2z8dm/. Results of a related study (delayed gratification) are statistically non-significant (10.17605/OSF.IO/WZYFC).

used in the experiment. Because each participant was asked two distance questions (one in each of the two languages), our final dataset includes 1140 observations (two per participant). On average the perceived distance between ‘today’ and ‘tomorrow’ was 0.60 (SD=0.35, N=570) and between ‘now’ and ‘later’ 0.54 (SD=0.31, N=570). Figure 2 presents the differences in the perceived distance by experimental condition.

On average, the perceived distance between the present and the future was 0.59 (SD=0.14, N=570) when participants were addressed in the strong FTR language and only 0.55 (SD=0.14, N=570) when participants were addressed in the weak FTR language ($p<0.05$).

In table 4, we report the results of OLS regression models predicting the perceived distance by whether questions were asked in the strong or weak FTR language and additional controls.

Model 1 includes a dummy variable for whether questions were asked in the strong (compared to the weak) FTR language and controls for the task (whether participants were asked to report the distance between ‘now’ and ‘later’ or the between ‘today’ and ‘tomorrow’). Model 2 additionally includes proficiency scores in the strong and weak FTR languages. Models 3 includes the gap in proficiency between the two languages and controls for the two languages that were used in the experiment. Model 4 includes the demographic characteristics of participants. Finally, model 5 is a person fixed effects regression model. Standard errors are clustered by person in all models.

As predicted in all models, being asked the distance question in a strong FTR language increased the perceived distance between the present and the future by 0.04 ($p<0.01$). Relatedly, in models 3 and 4, being more proficient in the strong FTR language increased the perceived distance between the present and the future (controlling for the effects of the specific languages used). Finally, even when person fixed effects regression models were estimated (Model 5), effects remained similar in magnitude and statistical significance. In other words, we observe the positive effects of being addressed in the strong FTR language on the perceived distance between the present and the future, even within person.

These within person findings suggest a causal effect of language on the perceived distance between the present and the future . They also provide stronger evidence for causality because individual unobserved differences between participants, attrition and selection biases are irrelevant (this is a within-person comparison; all the participants completed the questionnaires in both languages).

Discussion

The results of our first experiment suggest that people addressed in languages in which the present and the future are marked more distinctly tend to value future events less than people addressed in languages in which the present and the future are similarly marked. In addition, people more fluent in languages in which the present and the future are marked more distinctly (compared to their fluency in languages in which the present and the future are similarly marked) tend to value future events less, compared to people who are more fluent in languages in which the present and the future are similarly marked—regardless of the language in which they are being addressed. Finally, the effects of being addressed in languages in which the present and the future are marked more distinctly are stronger when participants are relatively more proficient in these languages, compared to their proficiency in languages in which the present and the future are similarly marked. The results of our second experiment show that when people are addressed in languages in which the present and the future are marked more distinctly they view the future as more distant compared to when they are addressed in languages in which the present and the future are similarly marked.

Our experimental designs enabled us to identify causality between the encoding of time in the language in which people are addressed and their tendency to view the future as distant and to discount future monetary payments. Taken together, the results suggest that the grammatical structure of the language in which one is addressed activates different perceptions of time, resulting in different time preferences and behaviors. In order to provide stronger evidence for causality, our second experiment is a within-participant experiment in which all participants were asked to mark the distance between the present and the future in two different languages that vary in the way they encode time. The findings support our research hypotheses.

Our studies have some limitations. FTR and modality (the strategies used to mark possible events) may correlate across languages. However, the EUROTYP project, which the FTR typology that we relied upon is based on, did not include any treatment of modality. We are not familiar with any quantitative typology of modality. Therefore, we cannot disentangle the effects of FTR from the effects of modality. Note, however, that it is unclear whether FTR and modality tend to correlate across languages. Interestingly, studies in psychology suggest that although thinking about the future and about counter-factual alternatives are distinct processes, they are related to each other cognitively, and similarly affect prediction, preference, and behavior (Trope and Liberman 2010).

Note also that the strong–weak FTR distinction used in this study is not confounded by any of the linguistic features that appear in the World Atlas of Language Structures. In this atlas, there are 56 linguistic features that have complete data for all the languages used in this study. None of these features could be used to divide the languages used here into the two groups we use, other than that of their future time reference. [6]

Our results suggest that the time-related schemas embedded in languages are easily and immediately activated; asking the same payment questions in a different language resulted in different time preferences for otherwise similar participants.

Languages both reflect and enforce time-related attitudes, preferences, and behaviors. The preferences of participants in the experiment were affected both by the encoding of time in the language in which they were addressed and by the encoding of time in the language in which they are more proficient. Thus, languages routinely and actively participate in enacting and maintaining schemas about time; whenever a language is spoken, the time preferences embedded in it are further reinforced, and behaviors follow accordingly.

---

[6] We are grateful to the anonymous reviewer who brought this evidence to our attention.

Bibliography

Campo, Francesco, Luca Nunziata, and Lorenzo Rocco. 2020. "Talking Business: New Evidence on How Language Shapes Economic Behaviour."

Carroll, Mary, Christiane Von Stutterheim, and Ralf Nüse. 2004. "The language and thought debate: a psycholinguistic approach." In Multidisciplinary Approaches to Language Production, ed. T Pechmann, C Habel, pp. 183–218. Berlin: Mouton de Gruyter

Chen, Keith M. 2013. "The Effect of Language on Economic Behavior: Evidence from Savings Rates, Health Behaviors, and Retirement Assets." American Economic Review, Vol 103, No. 2, pp 690-731.

Chen, Shimin, Henrik Cronqvist, S. X. Ni, and Frank Zhang. 2015. "Language and corporate decision-making." *SSRN Electronic Journal*.

Costa-Font, Joan, Paola Giuliano, and Berkay Ozcan. 2018. "The cultural origin of saving behavior." *PloS one* 13, no. 9: e0202290.

Dahl, Östen. "Tense and aspect systems." (1985).

Dahl, Östen. 2000. "The Grammar of Future Time Reference in European Languages." In *Tense and Aspebct in the Languages of Europe*, edited by Östen Dahl, 309–28. Berlin: Mouton de Gruyter.

Daniel J. Benjamin & James J. Choi & A. Joshua Strickland. 2010. "Social Identity and Preferences," *American Economic Review*, American Economic Association, vol. 100(4), pages 1913-28.

Davis, Lewis and Megan Reynolds. 2018. "Gendered language and the educational gender gap." *Economics Letters*, 168, 46-48.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman and Uwe Sunde. 2018. "Global Evidence on Economic Preferences." *The Quarterly Journal of Economics,* 133(4), 1645-1692.

Fasan, Marco, Giorgio Gotti, Tony Kang, and Yi Liu. 2016. "Language FTR and earnings management: International evidence." *Available at SSRN 2763922*.

Frederick, Shane, George Loewenstein, and Ted O'donoghue. 2002. "Time discounting and time preference: A critical review." *Journal of economic literature*, vol. 40, no. 2: 351-401.

Galor, Oded and Özak, Ömer and Sarid, Assaf. 2018. "Geographical Roots of the Coevolution of Cultural and Linguistic Traits." *NBER* Working Paper No. w25289. Available at SSRN: https://ssrn.com/abstract=3290424.

Galor, Oded, Ömer Özak, and Assaf Sarid. 2020. "Linguistic traits and human capital formation." In AEA Papers and Proceedings, vol. 110, pp. 309-13.

Gay, Victor, Estefania Santacreu-Vasut and Amir Shoham. 2013. "The grammatical origins of gender roles." In: *Working papers Berkeley Economic History Laboratory Paper Series*.

Grossman, Eitan. 2013. "The costs of ignoring language change." Diversity Linguistics Comment Blog, March 10. http://dlc.hypotheses.org/371

Hübner, Malte, and Gonzague Vannoorenberghe. 2015. "Patience and long-run growth." *Economics Letters* 137: 163-167.

Jakiela, Pamela and Owen Ozier. 2018. "Gendered Language". Policy Research working paper; no. WPS 8464. Washington, D.C. : World Bank Group. http://documents.worldbank.org/curated/en/405621528167411253/Gendered-language

Karapandza, Rasa. 2016. "Stock returns and future tense language in 10-K reports." *Journal of Banking & Finance* 71: 50-61.

Kim, Jaehyeon, Yongtae Kim, and Jian Zhou. 2017. "Languages and earnings management." *Journal of Accounting and Economics* 63, no. 2-3: 288-306.

Ladd, D. Robert, Seán G. Roberts, and Dan Dediu. 2015. "Correlational studies in typological and historical linguistics." *Annu. Rev. Linguist.* 1.1: 221-241.

Levinson, Stephen C. 2012. Foreword. In Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf, ed. JB Carroll, SC Levinson, P Lee, pp. vii–xxiii. Cambridge, MA: MIT Press. 2nd ed.

Li, Peggy, and Lila Gleitman. 2002. Turning the tables: language and spatial reasoning. Cognition 83 (3): 265–294.

Liang, Hao, Christopher Marquis, Luc Renneboog, and Sunny Li Sun. 2018. "Future-time framing: The effect of language on corporate future orientation." *Organization Science* 29, no. 6: 1093-1111.

Liberman, Mark. 2013. "High-altitude ejectives." Language Log Blog, June 14. https://languagelog.ldc.upenn.edu/nll/?p=4685

Lucy, John A. 1996. "The Scope of Linguistic Relativity: An Analysis and Review of Empirical Research." In Rethinking linguistic relativity, John J. Gumperz, and Stephen C. Levinson (eds.), 37–69 Cambridge: Cambridge University Press.

Mavisakalyan, Astghik, Yashar Tarverdi and Clas Weber. 2018. "Talking in the present, caring for the future: Language and environment." *Journal of Comparative Economics*, 46(4), 1370-1387.

Nuyts, Jan. "Epistemic Modality, Language, and conceptualization" *Amsterdam ua: Benjamins* (2000).

Prewitt-Freilino, Jennifer L., Andrew T. Caswell and Emmi K. Laakso. 2012. "The gendering of language: A comparison of gender equality in countries with gendered, natural gender, and genderless languages." Sex Roles, 66, 268 –281.

Palmer, Frank Robert. *Mood and modality*. Cambridge university press, 2001.

Pullum, Geoffrey K. 2012. "Keith Chen, Whorfian economist." Language Log Blog, Feb. 9. http://languagelog.ldc. upenn.edu/nll/?p53756

Roberts, Sean G, James Winters and Keith Chen. 2015. "Future Tense and Economic Decisions: Controlling for Cultural Evolution." *PLoS ONE* 10(7): e0132145. https://doi.org/10.1371/journal.pone.0132145

Sutter, Matthias, Silvia Angerer, Daniela Glätzle-Rützler and Philipp Lergetporer. 2018. “Language group differences in time preferences: Evidence from primary school children in a bilingual city.” *European Economic Review*, vol. 106, issue C, 21-34.

Thieroff, Rolf. "On the areal distribution of tense-aspect categories in Europe." *Empirical approaches to language typology* 6 (2000): 265-308.

Trope, Yaacov, and Nira Liberman. 2010. "Construal-level theory of psychological distance." *Psychological review* 117, no. 2: 440.

Whorf, Benjamin. 1956. Carroll, John B. (ed.), Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf, MIT Press.

Tables

Table 1: Attrition					
	Weak-FTR		Strong-FTR		Difference
	Obs	Proportion	Obs	Proportion	P-Value
Step					
Assignment to Treatment	3033		3156		.118
Answered Payment Questions	1980	.65	2213	.70	.000
Completed Survey	1822	.60	1982	.63	.027
Proficient in Both Languages	364	.12	353	.11	.316
Unique Users	310	.10	295	.09	.247
Consistent Time Preference	289	.10	276	.09	.285

Table 2: Sample Characteristics by Experimental Condition

	Weak-FTR	Strong-FTR
Language Pairs		
English-Dutch	0.13	0.11
English-German	0.21	0.23
English-Mandarin	0.15	0.17
French-Dutch	0.04	0.05
French-German	0.06	0.10
French-Mandarin	0.04	0.03
Hindi-Dutch	0.05	0.05
Hindi-German	0.08	0.04
Hindi-Mandarin	0.05	0.03
Spanish-Dutch	0.06	0.05
Spanish-German	0.10	0.08
Spanish-Mandarin	0.04	0.05
Payment reservation price		
$3.05	0.33	0.30
$3.25	0.11	0.11
$3.50	0.07	0.08
$3.75	0.06	0.04
$4.00	0.13	0.13
$4.50	0.06	0.08
$5.00	0.10	0.08
$5.50	0.02	0.02
$6.00	0.02	0.04
$7.00	0.03	0.05
None selected	0.06	0.09
Proficiency in the Addressing Language	7.45	8.15
	(1.13)	(1.09)
Strong-Weak Proficiency Gap	0.51	0.54
	(1.15)	(1.20)
Female	0.38	0.37
White/Caucasian	0.41	0.41
African American	0.02	0.02
Hispanic	0.05	0.06
Asian	0.48	0.47
Other Race	0.04	0.04
College	0.78	0.79
Strong Language Genus:		
Germanic	0.48	0.51
Indic	0.18	0.12
Romance	0.33	0.37
Weak Language Genus		
Germanic	0.28	0.28
Indic	0.72	0.72
N	289	276

The table reports group means. Standard errors are in parentheses.

Table 3: TOBIT Regression Models Predicting Lowest Accepted Delayed Payment					
	(1)	(2)	(3)	(4)	(5)
Asked in Strong FTR	0.504***	0.497***	0.276ⁱ	0.521***	0.515***
	(0.145)	(0.145)	(0.149)	(0.143)	(0.144)
Proficiency in the Addressing Language	-0.437***	-0.431***	-0.603***	-0.451***	-0.433***
	(0.063)	(0.063)	(0.073)	(0.066)	(0.066)
Strong Weak Proficiency Gap		0.118*	-0.189*	0.077	0.074
		(0.056)	(0.085)	(0.061)	(0.059)
Asked in Strong FTR X Strong Weak Proficiency Gap			0.607***		
			(0.130)		
Female				0.158	0.177
				(0.141)	(0.142)
African American				0.519	0.591
				(0.447)	(0.453)
Hispanic				-0.234	-0.233
				(0.316)	(0.312)
Asian				-0.580***	-0.644***
				(0.178)	(0.174)
Other				-0.036	-0.071
				(0.383)	(0.387)
College Graduate				-0.239	-0.266
				(0.174)	(0.175)
Language Pairs Dummies				Y	
Strong-FTR Genus Indic					-0.133
					(0.219)
Strong-FTR Genus Romance					-0.105
					(0.161)
Weak-FTR Genus Indic					-0.186
					(0.168)
Constant	6.744***	6.638***	8.086***	7.165***	7.272***
	(0.472)	(0.473)	(0.564)	(0.565)	(0.570)
Sigma	2.162***	2.148***	2.062***	1.949***	2.021***
	(0.186)	(0.184)	(0.177)	(0.168)	(0.175)
N	523	523	523	509	509
Pseudo R-square	0.032	0.035	0.049	0.063	0.054

ⁱ p<0.1 * p<0.05 ** p<0.01 ***p<0.001

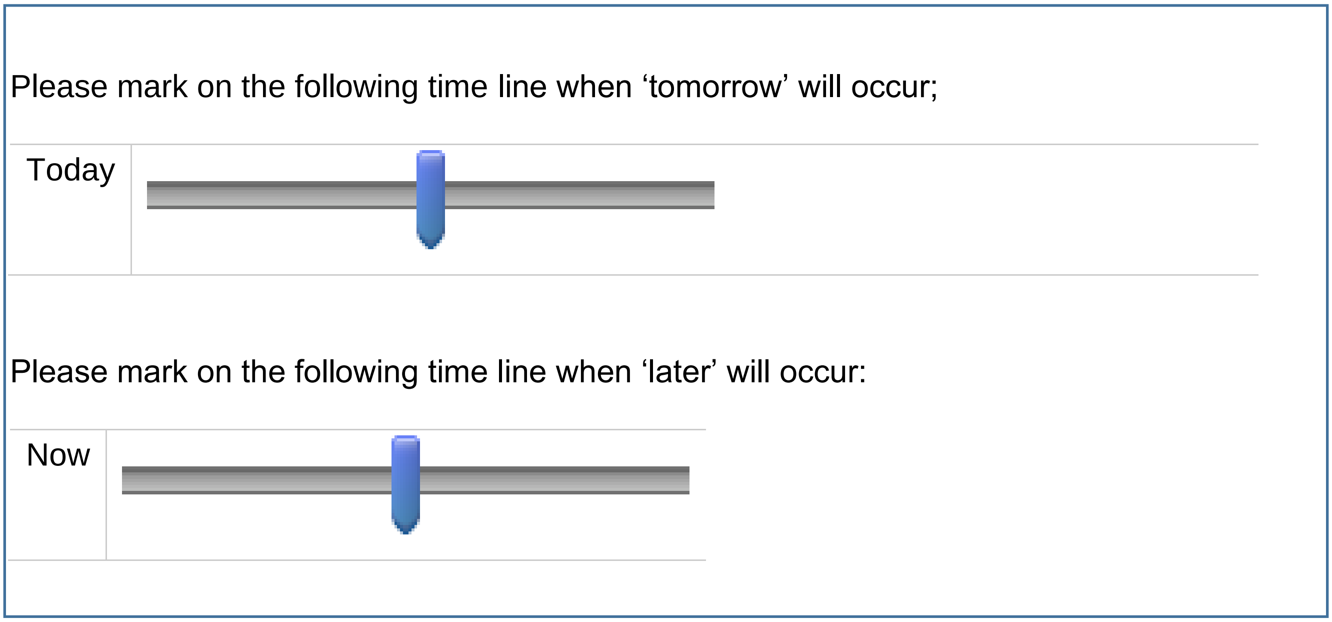
Table 4: OLS regression models predicting the perceived distance between the present and the future

	(1)	(2)	(3)	(4)	(5)
Asked in strong FTR	0.041***	0.041***	0.041***	0.041***	0.042**
	(0.012)	(0.012)	(0.012)	(0.012)	(0.018)
Task ('now' and later')	-0.060***	-0.060***	-0.060***	-0.060***	
	(0.012)	(0.012)	(0.012)	(0.012)	
Proficiency: strong FTR		-0.042	-0.084**	-0.071**	
		(0.028)	(0.035)	(0.036)	
Proficiency: weak FTR		-0.036*			
		(0.019)			
Proficiency gap :(strong-weak FTR)			0.037*	0.032	
			(0.021)	(0.021)	
Age				0.001	
				(0.001)	
Female				0.029	
				(0.033)	
Non-binary				-0.061	
				(0.129)	
Non-white				-0.023	
				(0.043)	
High school or less				0.055	
				(0.052)	
Constant	0.581***	1.263***	1.315***	1.102***	0.551***
	(0.016)	(0.292)	(0.314)	(0.326)	(0.009)
Language dummies					
			Y	Y	
Person Fixed effects					Y
N	1140	1140	1140	1140	1140

Standard errors in parentheses, clustered by person
* p<0.1, **p<0.05, ***p<0.01

Figures

Figure 1: sliders used in experiment 2



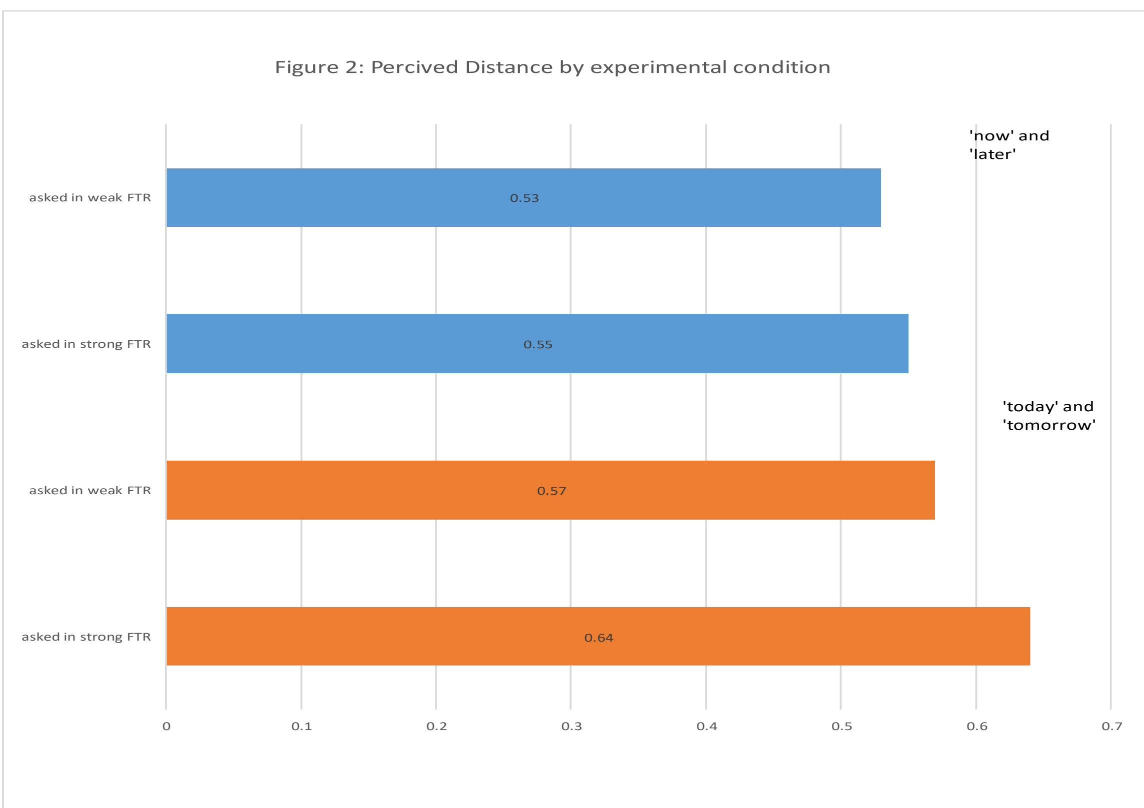


Figure 2: Percived Distance by experimental condition

SI Appendix Tables

Table S1: OLS Regression Models Predicting Lowest Accepted Delayed Payment

	(1)	(2)	(3)	(4)	(5)
Asked in Strong FTR	0.303***	0.298**	0.179*	0.295**	0.293**
	(0.091)	(0.091)	(0.094)	(0.095)	(0.092)
Proficiency in the Addressing Language	-0.252***	-0.248***	-0.343***	-0.232***	-0.219***
	(0.039)	(0.040)	(0.045)	(0.041)	(0.040)
Strong Weak Proficiency Gap		0.056	-0.122*	0.054	0.047
		(0.041)	(0.050)	(0.049)	(0.048)
Asked in Strong FTR X Strong Weak Proficiency Gap			0.347***		
			(0.088)		
Female				0.159ⁱ	0.177*
				(0.090)	(0.090)
African American				0.514ⁱ	0.571ⁱ
				(0.312)	(0.309)
Hispanic				-0.148	-0.154
				(0.194)	(0.188)
Asian				-0.288*	-0.334**
				(0.118)	(0.115)
Other				0.019	0.007
				(0.302)	(0.301)
College Graduate				-0.175	-0.200
				(0.132)	(0.132)
Language Pairs Dummies				Y	
Strong-FTR Genus Indic					0.109
					(0.153)
Strong-FTR Genus Romance					0.044
					(0.103)
Weak-FTR Genus Indic					-0.092
					(0.110)
Constant	5.731***	5.671***	6.478***	5.685***	5.697***
	(0.307)	(0.318)	(0.363)	(0.366)	(0.373)
N	523	523	523	509	509
Adjusted R-square	0.071	0.073	0.100	0.113	0.100

Robust standard errors in parentheses; ⁱ $p<0.1$ * $p<0.05$ ** $p<0.01$ *** $p<0.001$

Table S2: Ordered Logit Regression Models Predicting Lowest Accepted Delayed Payment

	(1)	(2)	(3)	(4)	(5)
Asked in Strong FTR	0.574***	0.639***	0.391**	0.705***	0.673***
	(0.159)	(0.156)	(0.167)	(0.165)	(0.158)
Proficiency in the Addressing Language	-0.588***	-0.596***	-0.814***	-0.645***	-0.616***
	(0.074)	(0.070)	(0.087)	(0.077)	(0.078)
Strong Weak Proficiency Gap		0.081	-0.274***	0.011	0.001
		(0.062)	(0.083)	(0.075)	(0.073)
Asked in Strong FTR X Strong Weak Proficiency Gap			0.719***		
			(0.144)		
Female				0.002	0.019
				(0.163)	(0.161)
African American				0.642	0.713 ⁱ
				(0.414)	(0.392)
Hispanic				0.328	0.369
				(0.359)	(0.366)
Asian				-0.631**	-0.661**
				(0.236)	(0.223)
Other				1.002 ⁱ	0.901 ⁱ
				(0.516)	(0.504)
College Graduate				-0.322	-0.325
				(0.211)	(0.210)
Language Pairs Dummies				Y	
Strong-FTR Genus Indic					-0.289
					(0.298)
Strong-FTR Genus Romance					-0.232
					(0.181)
Weak-FTR Genus Indic					-0.146
					(0.192)
cut1	-5.070***	-5.158***	-6.999***	-6.183***	-6.124***
	(0.583)	(0.550)	(0.701)	(0.709)	(0.734)
cut2	-4.511***	-4.620***	-6.441***	-5.608***	-5.556***
	(0.574)	(0.542)	(0.689)	(0.700)	(0.724)
cut3	-4.154***	-4.286***	-6.095***	-5.245***	-5.199***
	(0.567)	(0.536)	(0.682)	(0.693)	(0.717)
cut4	-3.919***	-4.070***	-5.872***	-5.014***	-4.972***
	(0.561)	(0.531)	(0.677)	(0.688)	(0.711)
cut5	-3.223***	-3.459***	-5.243***	-4.362***	-4.333***
	(0.545)	(0.518)	(0.661)	(0.677)	(0.698)
cut6	-2.757***	-3.082***	-4.858***	-3.970***	-3.950***
	(0.533)	(0.510)	(0.653)	(0.669)	(0.689)
cut7	-1.908***	-2.494***	-4.261***	-3.340***	-3.330***
	(0.529)	(0.509)	(0.651)	(0.666)	(0.681)
cut8	-1.664**	-2.355***	-4.120***	-3.207***	-3.198***
	(0.528)	(0.508)	(0.650)	(0.666)	(0.681)
cut9	-1.088*	-2.084***	-3.846***	-2.923***	-2.917***
	(0.535)	(0.510)	(0.650)	(0.668)	(0.681)
cut10		-1.575***	-3.330***	-2.390***	-2.388***
		(0.525)	(0.658)	(0.677)	(0.691)
N	523	565	565	551	551
Pseudo R-square	0.032	0.033	0.043	0.055	0.049

Robust standard errors in parentheses; ⁱ p<0.1 * p<0.05 ** p<0.01 ***p<0.001

Table S3: Balancing Tests (experiement 1)

	Asked in Weak FTR (1)	Asked in Strong FTR (2)	t-test p-value (1)-(2)
Strong FTR Language: English	0.484	0.507	0.588
	(0.029)	(0.030)	
Strong FTR Language: French	0.131	0.185	0.083*
	(0.020)	(0.023)	
Strong FTR Language: Hindi	0.183	0.123	0.047**
	(0.023)	(0.020)	
Strong FTR Language: Spanish	0.201	0.185	0.632
	(0.024)	(0.023)	
Weak FTR Language: Dutch	0.270	0.275	0.884
	(0.026)	(0.027)	
Weak FTR Language: German	0.453	0.446	0.856
	(0.029)	(0.030)	
Weak FTR Language: Mandarin	0.277	0.279	0.954
	(0.026)	(0.027)	
Strong FTR proficiency score	8.076	8.152	0.408
	(0.064)	(0.065)	
Weak FTR proficiency score	7.453	7.511	0.545
	(0.067)	(0.068)	
N	289	276	

Robust standard errors in parentheses. * p<0.1 ** p<0.05 ***p<0.01

Table S4: Tobit Regression Models Predicting Lowest Accepted Delayed Payment

	Excluding French	Excluding Hindi
	(1)	(2)
Asked in Strong FTR	0.544***	0.405**
	(0.161)	(0.150)
Proficiency in the Addressing Language	-0.390***	-0.409***
	(0.069)	(0.067)
Strong Weak Proficiency Gap	0.120*	0.080
	(0.061)	(0.057)
Constant	6.291***	6.599***
	(0.525)	(0.508)
Sigma	2.139***	1.948***
	(0.200)	(0.177)
N	443	443
Pseudo R-sq	0.031	0.031

Robust standard errors in parentheses. * p<0.05 ** p<0.01 ***p<0.001

Table S5: TOBIT Regression Models Predicting Lowest Accepted Delayed Payment, Excluding English					
	(1)	(2)	(3)	(4)	(5)
Asked in Strong FTR	0.441*	0.476*	0.333	0.455*	0.462*
	(0.215)	(0.215)	(0.211)	(0.209)	(0.21)
Proficiency in the Addressing Language	-0.611***	-0.613***	-0.813***	-0.558***	-0.521***
	(0.094)	(0.093)	(0.107)	(0.093)	(0.093)
Strong Weak Proficiency Gap		0.167¡	-0.228¡	0.153	0.152
		(0.094)	(0.133)	(0.093)	(0.094)
Asked in Strong FTR X Strong Weak Proficiency Gap			0.856***		
			(0.213)		
Female				0.338	0.389¡
				(0.213)	(0.217)
African American				1.051	1.108
				(0.775)	(0.79)
Hispanic				-0.634	-0.541
				(0.402)	(0.4)
Asian				-0.879***	-0.847**
				(0.261)	(0.26)
Other				0.395	0.328
				(0.487)	(0.495)
College Graduate				-0.586*	-0.581¡
				(0.293)	(0.297)
Language Pairs Dummies				Y	
Strong-FTR Genus Indic					0.026
					(0.253)
Weak-FTR Genus Indic					-0.557*
					(0.24)
Constant	7.939***	7.895***	9.510***	8.642***	8.293***
	(0.692)	(0.688)	(0.8)	(0.786)	(0.754)
/					
Sigma	2.481***	2.449***	2.299***	2.058***	2.173***
	(0.315)	(0.311)	(0.29)	(0.263)	(0.278)
N	260	260	260	254	254
Pseudo R-square	0.058	0.062	0.082	0.116	0.101

¡ p<0.1 * p<0.05 ** p<0.01 ***p<0.001

Table S6: Descriptive statistics (experiment 2)

	mean	sd	min	max	
Female	0.809	0.393			
Non binary	0.005	0.724			
High school or less	0.082	0.275			
Age	47.174	13.696	20		122
Proficiency: strong FTR	8.830	0.432	6		9
Proficiency: weak FTR	8.581	0.666	6		9
Proficiency gap (strong -weak FTR)	0.249	0.782	-3		3

N=570

Table S7: Language Pairs (experiment 2):

pair	Freq.	Percent
English Dutch	96	16.84
English German	103	18.07
English Mandarin	54	9.47
French Dutch	94	16.49
French German	70	12.28
French Mandarin	24	4.21
Spanish Dutch	77	13.51
Spanish German	52	9.12

N=570