Language Shapes Public Attitudes Toward Gender Equality

Efrén O. Pérez
Vanderbilt University

Margit Tavits
Washington University in St. Louis

Abstract

Can the way we speak affect the way we think about gender equality? Languages vary by how much they require speakers to attend to gender. *Genderless* tongues (e.g., Estonian) do not oblige speakers to designate an object's gender, while *gendered* tongues do (e.g., Russian). By neglecting to distinguish between male and female objects, we hypothesize that speakers of genderless tongues will express more liberalized attitudes toward gender equality. We support this claim with two survey experiments that randomly assigned the interview language to Estonian/Russian bilinguals, and a cross-national analysis of language effects on survey responses. Our evidence yields two main conclusions. First, speaking a genderless tongue reliably increases support for policies and efforts to combat gender imbalances. Second, language guides people's views about gender equality only when social norms about acceptable behavior are absent.

Many women across different nations still lag behind men in several domains (Blau, Brinton, Grusky 2006; OECD 2012), particularly in politics, where they are woefully under-represented and under-placed (Arriola and Johnson 2014; Hinojosa 2012; O'Brien 2015; O'Brien and Rickne 2016; Paxton et al. 2010). One line of investigation suggests that patriarchal attitudes and beliefs promote and maintain gender inequality (Epstein 2007; Inglehart and Norris 2003; Iversen and Rosenbluth 2010), with asymmetrical attitudes toward females affecting women's political representation (Huddy and Terkildsen 1993; Sanbonmatsu 2002) and economic opportunities (Fortin 2005), even in highly developed societies. Nevertheless, this research has difficulty explaining where these sentiments arise from in the first place and why they persist.

Part of the answer, we believe, has to do with the language one speaks. Languages vary by the degree to which they require speakers to attend to and encode gender (Boroditsky et al. 2003; Corbett 1991; Cubelli et al. 2011; Vigliocco et al. 2005). *Genderless* languages, such as Estonian or Finnish, do not require speakers to designate the gender of objects—even the word for "he" and "she" is the same in these tongues. In contrast, *gendered* languages, like Spanish and Russian, require speakers to differentiate genders and assign it to objects.[1] Spanish speakers, for instance, must mark the object "moon" as feminine by using the definite article *la*, as in *la luna*. Russian speakers face an even more demanding task. To say, for example, "a yellow moon shone in the sky" (желтая луна сияла в небе), they must designate the "moon," the color

---

[1] Languages can also fall somewhere in between these two extremes to the extent that they contain gender markings (Dryer and Haspelmath 2013). For example, English is only weakly gendered: although it uses gendered pronouns, it does not require speakers to assign gender to objects. According to the *World Atlas of Language Structures*, in 19 of the 141 countries covered, the dominant language is genderless, and in 39 it is gendered; the remaining languages contain varying degrees of gender markers.

"yellow," and the verb "shone" as feminine. Such grammatical rules make gender a highly salient category for speakers.

If language sets a frame of mind for how people think, then differences in gender markings across languages might partly account for individual differences in attitudes about gender equality. Cognitive psychologists have found that language subtly, but powerfully, affects human thinking (Boroditsky 2001; Boroditsky and Gaby 2010; Boroditsky et al. 2003; Fuhrman et al. 2011; Slobin 1996), which opens the door to possible language effects on political opinions. We theorize that speaking a genderless tongue helps to promote greater perceived equity between men and women by neglecting to formally distinguish between male and female objects. Speakers of such languages are likely to find it harder to perceive a "natural" asymmetry between the sexes. This, in turn, affects people's opinions about gender equality, where speakers of genderless tongues express more support for political efforts seeking to rectify gender imbalances.

In addition to outlining *how* language might impact attitudes toward gender equality, we also theorize about *when* its influence is less likely. Some research suggests that language has a stronger grip on thinking if the domain in question is more abstract—that is, if sensory information is constrained or inconclusive (Boroditsky et al. 2003; Echterhoff 2008; Phillips and Boroditsky 2003). We adapt this insight to politics and predict that language is less likely to matter when strong social norms exist about appropriate behavior, thus leaving less room for language to further affect one's opinions about gender inequality.

We test our claims with two original survey experiments that randomly assigned the interview language to bilingual adults in Estonia who speak equally well a gendered (Russian) and a genderless (Estonian) language. This parsimonious design lets us clearly identify the effect

of speaking a gendered or genderless language on people's views about gender equality (Dunning 2016; Green 2004). We then validate and extend our experimental findings on language by undertaking a cross-national analysis of survey data.

Consistent with our theory, we find that interviewing in a genderless tongue meaningfully affects people's attitudes about gender equality. Respondents who were assigned to interview in Estonian express more liberalized attitudes compared to respondents who were assigned to interview in Russian: a result that is also borne out by our cross-national examination of survey data. Critically, these language effects emerge only when a policy or initiative touches on a domain where normatively appropriate behavior is unclear, with a placebo test further showing that such language effects are confined to domains that clearly evoke gender.

Our findings provide new insights into the persistence of gender inequality – one of the most fundamental social divides (Epstein 2007). We reveal that simple quirks in language can impose important constraints on one's view of gender roles and affect how mentally receptive people can be to efforts toward eliminating gender inequalities.

More generally, our study highlights the potential benefits of borrowing insights from cognitive science to better understand key political questions. Prior research in political science has studied language use to better understand ethnicity and ethnic relations (e.g., Adida et al. 2016; Garcia Bedolla 2005; Laitin 1998; Laitin, Moortgat, and Robinson 2012; Posner 2005). Recent research also shows that language skills and non-native language use affect attitudes toward immigrants and their integration (Hopkins 2014, 2015; Hopkins et al. 2014; Sobolewska et al. 2016). Closer to our own study, Laitin's (1977) pioneering analysis in Somalia shows that the language we speak affects how we solve conflicts, while Garcia (2009) reveals the influence of interview language on survey responses among U.S. Latinos. We break new ground in this

research field by establishing that the language we speak can mold our construction of political reality. This is a novel way to think about language's impact on politics and calls for political scientists to further explore how other linguistic differences – beyond grammatical gender – might shape mass opinion, pushing forward research in various areas. For example, better understanding how language shapes thought can advance research in political and social psychology by broadening our understandings of opinion formation and change (Taber and Young 2013; Valentino and Nardis 2013). Greater knowledge of language effects on people's attitudes, behavior, and preferences can also shed new light on cross-country variation in policy choices and outcomes (Dalton 2014), relations between ethnic groups (Horowitz 1985), international interactions (Bueno de Mesquita and Smith 2012), and many other phenomena. Finally, scholars can advance research on political culture by using linguistic nuances as a way to systematically analyze and quantify cultural effects (Almond and Verba 1963).

Theory: How Language Affects Attitudes about Gender Parity

The notion that language shapes thinking is often associated with Benjamin Whorf, who proposed that people's thoughts are completely determined by the tongue that they speak (Boroditsky et al. 2003). Although many researchers have found Whorf's hypothesis untenable and unfalsifiable, a new generation of scholars has rehabilitated his claim into weaker but still important and testable versions (cf. Boroditsky 2001; Fuhrman et al. 2011; Marian and Neisser 2000; see also Kay and Kempton 1984; Lakoff 1987). Key to this refashioning is Slobin's (1996) notion of "thinking for speaking," which argues that languages vary in their grammatical organization, thereby obliging speakers to focus on different aspects of their experience when using a given tongue. For example, in order to say "the child ate the ice cream" in English, one must include the past tense. But to utter the same phrase in Russian, one must include the past

tense, note the child's gender, and state whether the child ate all or some of the ice-cream. Thinking for speaking is therefore very distinct for Russian and English speakers.

Slobin's framework implies that varied language speakers may be biased toward attending to those aspects of the world that their tongue demands (Boroditsky 2001). If speaking a language requires one to make certain distinctions between objects (e.g., different colors, gender, and time orderings), then the speaker may take for granted that these categories actually exist in the world and are relevant (Boroditsky 2001; Boroditsky et al. 2003; Boroditsky and Gaby 2010; Danziger and Ward 2010; Fuhrman et al. 2011; Hunt and Agnoli 1991). Hence, grammatical differences between tongues provide a mechanism through which language shapes thought, where a grammatical habit of speech leads to a habit of mind.[2]

Most important for our purposes are the grammatical distinctions that languages make with respect to gender. Gendered languages have grammatical structures that train speakers to focus on gender, while non-gendered tongues do not force speakers to make these distinctions. Careful research reveals that speakers of gendered languages are more keenly aware of gender differences: they are more likely to categorize the world in feminine and masculine terms and to project gender features onto objects and individuals (Boroditsky et al. 2003; Cubelli et al. 2011; Konishi 1993; Phillips and Boroditsky 2003; Sera et al. 1994).[3] Gendered language speakers are also more likely to attain their own gender identity sooner than speakers of less gendered tongues

---

[2] The argument here is not about vocabulary or other surface differences between languages. Rather, we are concerned with fundamental concepts, like gender, that have been made part of grammar (Lakoff 1987). Such grammatical concepts "are used *in* thought, not just *as objects of* thought," and they are used automatically and unconsciously, thus producing a significant "impact on how we understand everyday life" (Lakoff 1987, 335).

[3] For example, when Russian speakers were asked to personify days of the week, they generally personified grammatically masculine days as males and grammatically feminine days as females (Jakobson 1966). Relatedly, young Spanish speakers generally rated object photos as masculine/feminine according to their grammatical gender (Sera et al. 1994).

(Guiora et al. 1982).[4] Accumulated work also reveals that language effects like these arise from structural (i.e., grammatical) differences between tongues,[5] and that they do so on both linguistic *and* non-linguistic tasks, which indicates that language can impact cognition with little to no verbalization (Fuhrman et al. 2011; Fausey and Boroditsky 2011).

Building on these psychological insights, we argue that people's views about gender equality are affected, in part, by how gender is dealt with by their language on an everyday basis. Gendered tongues require people to explicitly distinguish between males and females, which should strongly activate gender as a concept in memory and make it more mentally accessible (Lodge and Taber 2013). This is crucial, since leading models of survey response suggest that rather than having pre-formed opinions on all matters, individuals construct their opinions on the basis of those considerations "at the top of the head" (Zaller 1992; Tourangeau et al. 2000). Speakers of gendered languages are therefore sensitized to the feminine or masculine qualities of individuals or objects, which is why they are likely to perceive gender differences as more salient and the roles of men and women as more distinct and divided.

We expect the opposite among speakers of genderless tongues. These languages minimize gender's salience as a significant category by not requiring its speakers to make distinctions on this basis. Therefore, speakers of genderless languages are less likely to perceive gender differences and more likely to see the roles of men and women as similar. If our reasoning is correct, then this variation in gender's salience should be reflected in people's

---

[4] Guiora et al. (1982) studied children, ages 16-42 months, who spoke tongues varying by their gendered-ness: Hebrew (highly gendered), English (medium gendered), and Finnish (genderless). By 28-30 months of age, 50% of Hebrew speakers expressed gender identification compared to 21% of English speakers and 0% of Finnish speakers.
[5] For example, researchers have taught individuals fictitious languages that are completely stripped of any cultural or other contextual information and found that that these tongues still affect individual behavior in predicted ways (Boroditsky et al. 2003).

opinions about gender equality. Specifically, we hypothesize (H1) that speakers of genderless tongues will express less stereotypical and more liberalized opinions about women's place in politics and society. We also expect to find greater support for *political efforts* to combat gender inequality among speakers of a genderless rather than gendered tongue. The implication here is that a tongue's genderedness plays a crucial role in shaping public opinion.

While our argument so far strongly suggests that language is likely to affect one's views of gender equality, it is also reasonable to expect that language may not always matter. Prior work reveals that language effects are more likely in situations lacking a universal understanding. These include domains that are abstract, rather than concrete; and domains where direct sensory information is inconclusive, rather than dispositive (Boroditsky 2001; Echterhoff 2008; Winawer et al. 2007). Clear and widely shared cues, then, provide additional, non-linguistic information that can drown out language effects (Boroditsky 2000, 2001).

Adapting these insights to politics, we theorize that language will fail to affect opinions about gender equality when stronger norms exist about socially acceptable behaviors or beliefs (H2). Such norms equip people with more information about a topic (i.e., provide a common understanding), thus leaving less space for language to affect people's judgments. When social norms are weak or absent, language will shape one's political thinking by filling these information gaps. For example, we might expect language to have a trivial impact on whether one thinks it is sometimes justified for a woman to hit her husband, given a social norm against aggression, regardless of whether its target is female or male. But we *would* expect language to influence whether more women should be recruited to top government posts because it is less clear what the social consensus is on such a matter.

Research Design

We test our claims across three independent studies. Study 1 consists of an original survey experiment we administered in Estonia from May 26 to June 12, 2014. Here, twelve hundred (N=1,200) Estonian-Russian bilingual adults were randomly assigned to interview in Estonian (i.e., a genderless tongue) or Russian (i.e., a gendered tongue).[6] Study 2 is a smaller experiment (N=262) that uses the same design and replicates and extends our initial findings from Study 1. We conducted this second study from March 22 to April 10, 2016. These experiments provide a compelling and straightforward between-subjects design for identifying language effects (Dunning 2016; Green 2004).

Estonia is an ideal setting for testing our hypothesis for several reasons. First, it possesses a sizeable population that is equally proficient in a gendered (Russian) and a genderless (Estonian) language: about 69% of the population identify Estonian and 29% Russian as their first language. Roughly 44% of the former group and 36% of the latter speak the other language well enough to qualify as bilingual according to our definition.[7] Second, prior research shows that in terms of political opinions and values, Estonians and Russians in Estonia have more in common with each other than with any other group outside Estonia (Lauristin and Vihalemm 1997; Maimone 2004). Russians in Estonia do not express more traditional or conservative

---

[6] Given the novelty of this first study, we wanted to ensure that our survey experiment had enough statistical power to detect non-trivial opinion differences. Mean differences with Cohen's $d = 0.20$ and two-tailed $p < 0.05$ require N = 1,054. Our sample (N = 1,200) can therefore unearth meaningful language effects, in either direction, if they in fact exist.

[7] While its current official language is Estonian, the country was part of the USSR until 1991, which made Russian a prominent tongue, with most Estonian speakers acquiring at least a working knowledge of Russian and many becoming proficient. A large-scale immigration of Russians before 1991 also created a sizeable Russian-speaking population that acquired proficiency in the official language – Estonian. Segregated communities, inter-marriages, and schools offering general education (equivalent to K-12) in either Estonian or Russian further increased the bilingual population.

values than Estonians do. We demonstrate this with our own placebo test in Study 2. This makes

Estonia a uniquely ideal setting to cleanly identify language effects in public opinion.[8]

Study 3 further alleviates any concerns over the generalizability of the findings from our

Estonian experiments. This last study is a cross-national analysis of survey data from the World

Values Survey (WVS) designed to appraise the external validity of our experimental results. We

now describe in further detail each study and its relevant results.

Study 1: Survey Experiment with Estonian-Russian Bilinguals

Eligible bilinguals were identified via self-rated skill in Estonian and Russian. Bilingual

respondents who said they "can understand, speak, and write" or are "fluent" in both languages

were randomly assigned to interview in Estonian or Russian. Online Appendix (OA) section

OA.1 provides further details about the identification of bilinguals and the survey protocol,

including our language manipulation and English translations of question wording.

Thirty-eight percent (38%) of our sample consists of bilinguals whose first language is

Russian, and who, on average, learned Estonian at the age of fifteen. In turn, those bilinguals in

our sample (62%) whose first language is Estonian learned Russian, on average, at the age of

fourteen. Finally, sixty-one percent (61%) of our respondents are female, with a median age of

52 and secondary level of education. We collected information on these and a few other

---

[8] One might argue that Finnish and Swedish speakers in Finland offer another possible research
site for this study. While Finnish (which is very close to Estonian) is clearly a genderless
language, the status of Swedish as even weakly gendered language is questionable. While
Swedish has two grammatical genders, these are not masculine and feminine but "neuter" and
"common gender." Furthermore, Sweden has been using a gender-neutral pronoun ("hen") since
2012, and formally adopted it in 2015 (Senden et al. 2015). This does not allow for a clean
design based on grammatical gender.

demographic characteristics of each respondent pre-treatment. These variables are balanced

across our experimental conditions (see OA.2).[9]

Attitudinal Measures of Gender Equality

Post-treatment, respondents answered several items related to their perceptions of women and

their role in society and politics, which we describe below (full item wording is in OA.1). We

constructed these items on the basis of prior research on attitudes about gender, and several of

them are adapted from well-established national or international surveys such as the General

Social Survey and the Americas Barometer. Our items are designed to tap attitudes about gender

roles and gender bias in several ways, including: (a) those expressed in gender stereotypes,

which foster unequal perceptions of men and women (Bauer 2015; Dolan 2014; Koch 2002); (b)

preferences over women's participation in politics and political leadership roles, where females

are substantially underrepresented (O'Brien 2015; O'Brien and Rickne 2016); and (c) support for

policies aimed at changing norms that promote gender equality. For our policy items, we picked

issues that are salient and topical in Estonian public discourse. More specifically, we included

the following questions about women and politics, which serve as our dependent variables.

Emotional women and *Emotional men* are seven-point scale ratings of how emotional

(keyed as 7) vs. rational (keyed as 1) bilinguals believe men and women to be, with the item

order randomized. We use these ratings in two ways. First, we analyze them individually and in

their original format (variable names *Emotional women: single rating* and *Emotional men: single

rating*). Second, we difference these ratings to create a measure ranging from -6 to 6, where

---

[9] By design, our respondents are a subset (but not a random sample) of highly bilingual Estonian adults who were randomly assigned to interview in Estonian or Russian in a phone survey. We did not calculate sample weights since we are only interested in estimating the causal effect of our manipulation in this research setting.

positive values indicate greater stereotypical belief in women as emotional (relative to men) (variable name *Emotional women: relative rating*) (cf. Kinder and Kam 2009).

Paternity leave queried bilinguals about whether they agreed (keyed as 1) or disagreed (keyed as 0) with a proposed change in family leave policy that would allow a father to stay home, while the mother can return to work as soon as she is able to. (The current legislation allows the father to stay home only after the baby is at least 70 days old.)

Female Defense Minister asked bilinguals "If the party that you normally like nominated a generally well-qualified woman to be Minister of Defense, would you support that choice?," with support coded as "1" and opposition as "0."

Female political recruitment and *Female candidates: men's expense* are items that bilinguals answered in random order. The former asked whether one strongly disagreed (4), somewhat disagreed (3), somewhat agreed (2), or strongly agreed (1) that women should be recruited to "top-level government positions." Using the same response scale, bilinguals also reported the degree to which political parties should be required "to reserve some space on their lists of candidates for women, even if they have to exclude some men." These items are left in their raw metric and coded so that higher values reflect greater agreement with a statement.

Finally, *Buy sex, Child suffers, Hit husband,* and *Men better* are agree/disagree items that allow a test of our proposed boundary condition: i.e., language effects are less likely when clearer social norms suggest acceptable behavior. *Buy sex* first informed bilinguals that "Under the current law, 'pimping' is a criminal act but buying sex is not," then asked them whether they agreed/disagreed that "buying sex should also be made illegal." *Child suffers* asked bilinguals whether they agreed/disagreed that "a preschool child is likely to suffer if his or her mother works." *Hit husband* queried about whether "it is sometimes justified for a woman to hit her

husband," while *Men better* asked whether "on the whole, men make better political leaders than women do." We see these items as strongly implying appropriate behaviors: e.g., "buying sex is wrong," "mothers should stay home with toddlers," etc. For all of these items, agreement is coded as "1" and disagreement as "0."

Results from Study 1

We first examine whether interviewing in Estonian (genderless tongue) affects how much asymmetry people perceive between men and women in terms of gender stereotypes. Accordingly, we estimate the degree to which Estonian interviewees see less stereotypical differences between men and women, where the dependent variable is the differenced ratings of females and males as more emotional than rational (variable name *Emotional women: relative rating*). On this relative measure, higher values reflect individual views of women as more emotional than men, with the scale midpoint indicating no sensed difference between men and women on this stereotypic trait.

We find that interviewing in Estonian significantly reduces how emotional women are seen relative to men. The mean value for respondents interviewing in Russian is 1.34 (CI: 1.17, 1.52) and in Estonian 1.14 (CI: 0.98, 1.30), for a difference in means: $t = 0.2$, $p = 0.09$, two-tailed test. We also performed a regression analysis, presented in the first column of Table 1 to confirm this result.[10] These results support hypothesis 1. After using randomization to hold constant all other (un)observed differences between respondents interviewing in Estonian vs. Russian, the former are still less likely to perceive women as more stereotypically emotional than men.

---

[10] Given our experimental design, we prefer to estimate our treatment effects by regressing each dependent variable on the randomly assigned language of interview. However, Table OA.3.1 shows that these estimated treatment effects remain largely intact if we adjust them for individual differences in age, education, and gender.

We now look at whether this cognitive mechanism travels to the political domain and affects people's opinions about gender equality by estimating the effect of interviewing in Estonian on our other dependent variables: *Paternity leave*, *Female Defense Minister*, and *Female political recruitment*. We reason that if, in fact, a genderless tongue leads its speakers to perceive less asymmetry between men and women, then respondents interviewing in Estonian should be more supportive of paternity leave, because men and women can both care for a child. Similarly, respondents interviewing in Estonian should be more likely to back a female nominee for Defense Minister, since a qualified woman can just as well fulfill a political post traditionally occupied by men; and, Estonian interviewees should endorse greater female political recruitment at a higher level, because capable women, like men, can also be effective politicians.

As the last three columns of Table 1 indicate, this is indeed what we find. Given the non-linear nature of these estimates, we delve more deeply into the substance of these results by translating the raw coefficients into predicted probabilities that we present graphically in Figure 1. Panel A (Figure 1) shows the shift in the probability of supporting changes in paternity leave policy. Among respondents assigned to interview in Russian, the probability of supporting this policy change is 35%. But if a person is assigned to interview in Estonian, the probability of endorsing this proposal climbs reliably by 8 points to 43% (first difference (FD) = 0.08, CI: 0.02 to 0.14). Thus, simply by interviewing in a genderless tongue, respondents are significantly more supportive of a parental leave policy that is more gender balanced.

Turning to panel B, when asked whether respondents would endorse their party's nomination of a female for the position of Defense Minister—a post typically occupied by males—the probability of supporting this nomination among those assigned to interview in Russian is 65%. But when assigned to report a response to the same question in a genderless

tongue, the probability of endorsing a female's nomination for Defense Minister increases to 73%, which is a reliable 8-point shift in support (FD = 0.08, CI: 0.02, 0.13). Hence, interviewing in a genderless tongue also makes a female nominee for Defense Minister discernibly more palatable to respondents.

A comparable pattern emerges when we consider increasing the profile of women in the higher echelons of government more generally. As panel C illustrates, respondents interviewing in Russian have a 23% probability of strongly agreeing with greater efforts to recruit females to top government positions. But for respondents interviewing in Estonian, the probability of strongly agreeing with the same proposal climbs to 28%, for a reliable 5-point shift in support (FD = 0.05, CI: 0.01, 0.09).[11]

Overall, we deem these results strong for two reasons. First, social psychologists have shown that a person's native language can impact their thinking in other tongues (Phillips and Boroditsky 2003). Since our bilingual respondents acquired one of their languages before the other one, it should be especially hard for us to uncover language effects like the ones we have found. Second, unlike prior laboratory studies on language, we unearth our language effects in a large and heterogeneous survey sample where the treatment was administered by phone. The fact that we observe the influence of language on political thinking in this new research setting enhances the external validity of prior language effects research (Campbell and Stanley 1963), thereby bolstering the claim that language affects human cognition.

Robustness

---

[11] In OA.4 we further demonstrate that the effect of speaking a genderless tongue stems from de-emphasizing distinctions between women and men (i.e., gender-blindness), rather than promoting females or devaluing males (i.e., pro-female bias).

To further bolster our conclusions, we assessed the robustness of our results in three ways (see OA.3). First, we re-estimate our treatment effects by including subjects' age, education, and gender as covariates) in our regression analysis. Including these pre-treatment variables yields treatment effects that are similar to those reported in the main analysis (Table OA.3.1). Second, we demonstrate that our estimated language effects are unaltered if we adjust them for subject's preferred interview language (Table OA.3.2). This further suggests that our language effects are situational, arising from random assignment to interview in a certain language, independent of any influence brought to bear by subjects' inclination to interview in a specific tongue. Finally, since all of our interviewers were female (see SI.1 for more on the interview protocol), we tested for the possibility that our interviewers lead male (female) subjects to express stronger anti-female opinions because they feel that women "should not" be working in visibly public roles, like interviewing survey subjects. If this is true, then subjects' gender should consistently moderate the effect of interview language on opinions. We find that it does not (Table OA.3.3).

Boundary Condition: When Are Language Effects Less Likely?

We now take a further step and probe whether there are cases when the language effects just described do not occur. To this end, we test our second hypothesis according to which language effects are less likely to emerge when strong social norms exist about appropriate behavior. Specifically, we analyze the effect of genderless language on four items that we believe are attended by strong social norms: *Buy sex*, *Child suffers*, *Hit husband*, and *Men better*. If our reasoning is correct, then we should observe negligible differences between the opinions of Estonian and Russian interviewees on each of these items. Figure 2 depicts the relevant results (Table OA.5 contains the raw estimates). As expected, across each of these items, respondents interviewing in Estonian are just as likely as those interviewing in Russian to agree that "buying

15

sex should be made illegal" (panel A), that a "child suffers if the mother works" (panel B), that "it is sometimes justified for a women to hit her husband" (panel C), and that "men are better political leaders than women" (panel D). In other words, the interview language has no influence on respondents' judgment if the domain in question is attended by clearer social norms about appropriate behavior. This suggests that while language can affect some of what people think about gender inequality, it does not influence all of their views on this broad topic.[12]

Study 2: Replication Experiment with Estonian-Russian Bilinguals

In order to bolster faith in Study 1's results, we conducted a second experiment with three goals in mind. First, we sought to replicate some of the effects from Study 1 in order to be more confident that these results were not flukes. Second, while our failure to find any language effects in the presence of stronger social norms is highly suggestive of a boundary condition, our inference here rests on null findings. Study 2 manipulates these norms directly, allowing for a stronger test of this aspect of our theory. Finally, Study 2 includes a placebo test to demonstrate that language fails to affect opinions that do not directly evoke gender distinctions.

Within resource constraints, we again partnered with the same survey firm from Study 1 to run a new experiment (N=262). Our smaller sample here comes with a reduction in statistical power, from Study 1's bountiful 0.90 to a more modest 0.60 in this new study. This decrease is partly offset by the results of Study 1, which yield directional hypotheses that we test here with

---

[12] These null results do not appear to stem from poorly designed survey questions. The bottom panel of Table OA.3.3 shows that these items are associated with individual differences in age, education, and gender in theoretically intuitive ways.

one-tailed significance tests. At this power level, then, detecting mean differences with Cohen's

$d = 0.20$ and one-tailed $p < 0.10$ requires N = 238.[13]

The setting, recruitment, and type of participant for Study 2 were the same as in Study

1.[14] Our question battery re-administered our *Paternity leave*, *Female Defense Minister*, and

Female political recruitment items, thus ensuring a fresh test of Study 1's core results. To this

slate, we added a question-wording manipulation that directly tests our social norms hypothesis.

This experiment contained two conditions. In the first, subjects were asked their degree of

agreement in "calling on party leaders to encourage more women to run for office" (*Run for*

office).[15] This item is similar in spirit to the items where we have so far found effects, and we

expect to find one here. The second condition is identical in wording to the first, except that it

concludes by describing this effort as "a proposal that about 80% of the people in Estonia favor."

If our reasoning is correct, then any language difference that we observe in the first condition

should decline, if not vanish, in this second condition, which makes a social norm salient.

Finally, as a placebo, we asked respondents to indicate how justifiable they think suicide

is, with 1 being "never justifiable" and 10 being "always justifiable." Since gender is not

(in-)directly referenced by this item, we do not expect language to matter here.[16] This item also

lets us probe an alternative explanation for our language effects: that speaking a language

activates ideological thinking. For example, one might argue that respondents assigned to speak

---

[13] This reduced power level means the odds of failing to reject the null hypothesis when it is *false* grow from 1 in 10 times (Study 1) to 2 in 5 times (Study 2), which works against uncovering meaningful effects in this new sample.

[14] See OA.1for the interviewer script that we used for the treatment.

[15] The response options were (1) strongly disagree, (2) somewhat agree, (3) somewhat agree, and (4) strongly agree. See OA.1 for precise question wording for the new items in Study 2.

[16] As Tables OA.6.1 and OA.6.2 show, we again find evidence suggesting an effective randomization and a lack of large and systematic imbalances in pre-treatment covariates.

Estonian support gender equality more, not because of the grammar distinction we propose, but because speaking Estonian primes respondents to think in more socially liberal terms. If this is the case, then respondents interviewing in Estonian should also be more likely to find suicide justifiable – a position that is opposite to the religious-conservative stance against suicide.

Results from Study 2

Table 2, under the label "Replication," reports Study 2's results. The evidence there is generally consistent with our main hypothesis: being assigned to interview in Estonian increases support for policies and efforts to address gender inequality. In particular, assignment to interview in Estonian significantly boosts support for a more flexible paternity leave policy (0.41, $p<0.01$, one-tailed); reliably increases respondents' backing of a female defense minister (0.26, $p<0.08$, one-tailed); and also appears to bump up support for greater female political recruitment, though this last effect is unreliable (0.01, $p<0.48$, one-tailed). Crucially, this general trend is robust to an alternative configuration of these three indicators. If we combine these items into a counter variable where a score of "1" is assigned to each liberal response that respondents give to each of these three questions, we find that interviewing in Estonian still positively and reliably boosts scores on this omnibus measure (0.42, $p<0.01$, one-tailed).[17] We interpret this evidence as indicating that the language effects we uncovered in Study 1 are unlikely to have been outliers.

The middle columns in Table 2, under the label "Social Norms Experiment," report the results of the question-wording manipulation we developed to directly test our social norms proposition. There we see that in the absence of any normative information, those who are

---

[17] The correlations between these indicators are positive and modest (average $\rho = 0.14$), but their magnitude is likely larger in the absence of measurement error, which is present here. For this exercise, the item *Female political recruitment* was dichotomized to align with the structure of the other two items.

assigned to interview in Estonian are, just as we predict, reliably more supportive of "calling on party leaders to encourage more women to run for office" (0.36, $p<0.05$, one-tailed). However, in the presence of normative information—that is, when respondents are told that 80% of Estonians favor this proposal—the effect of interviewing in Estonian drops in size and significance (0.22, $p>0.16$, one-tailed). Thus, salient social norms undercut the impact that language would have had on people's opinions about gender equality.[18]

The last column in Table 2, labeled "Placebo Test," contains the results for our item gauging opinions toward suicide. There we can see that our language effect is not insensitive to topic. Unlike those items where gender is clearly implicated, we find that here, interviewing in Estonian has no reliable influence on whether one deems suicide to be justifiable (0.15, $p>0.30$, one-tailed)—a domain we hardly see as evoking gender distinctions. Moreover, this test suggests that speaking Estonian does not simply prompt people to give more liberal answers.

All in all, Study 2's results corroborate and extend those in Study 1. We find that speaking a genderless language causes more support for gender equality; that this effect is dampened by social norms; and that this language effect arises only in domains that evoke gender distinctions. Taken together, these results strongly suggest that differences in grammatical gender across languages cause shifts in speakers' attitudes toward gender equality.

Study 3: Cross-National Evidence of Language Effects

We designed our experiments to detect language's causal effects, thereby achieving a high degree of internal validity. But this also restricted us to one specific context, which leads our experiments to attain a relatively low degree of external validity. The goal of Study 3, then, was to explore whether genderless languages are associated with more gender-balanced attitudes in

---

[18] The difference between these effects is statistically significant.

contexts beyond Estonia (cf. Hicks et al. 2015; Prewitt-Freilino et al. 2009; Santacreu-Vasut et al. 2013; Santacreu-Vasut et al. 2014). To this end, we used the World Values Survey (WVS), waves 3-6 (1995-2014),[19] to conduct a cross-national analysis of language effects that spans about 90 countries and, depending on the model, up to 170,000 individuals. These extensive survey data furnish us an unrivaled opportunity to replicate our main experimental findings in a global context.

We rely on the following four items to reassess our experimental results: (1) *Women jobs* is a binary variable coded "1" if the respondent disagrees that "when jobs are scarce, men should have more right to a job than women," and "0" otherwise; (2) *Women political leaders* inquired whether respondents strongly agreed (4), agreed (3), disagreed (2) or strongly disagreed (1) with the statement that "on the whole, men make better political leaders than women do;" (3) *University for girls* uses a similar 4-point scale to measure respondent's agreement with the statement that "university more important for a boy than for a girl;" (4) *Women business executives*, again, uses a similar 4-point scale to measure respondent's agreement with the statement that "on the whole, men make better business executives than women do."

These items parallel the main outcomes in Studies 1 and 2 and therefore serve as appropriate measures to validate our experimental results.[20] For all of these variables, higher values indicate more gender-equal responses. Since responses to these items are highly correlated ($\alpha = 0.74$), we also scale them to produce the variable *Gender equal attitudes*, where

---

[19] Earlier waves of WVS do not include information about the language spoken at home, which we need to perform our analysis of language effects.

[20] Furthermore, the first three items are included in all of the WVS waves covered in our study; the last item is included in waves 5 and 6. Collectively, these are the gender items that have the best coverage across the different waves of the WVS.

higher values indicate more gender-balanced responses. This index serves as the fifth dependent variable in the analyses reported below.

We rely on the question "What language do you normally speak at home?" to identify respondent's language. In order to measure the genderedness of the language spoken at home, we utilize data from the *World Atlas of Language Structures* (WALS) (Dryer and Haspelmath 2013), the most comprehensive data source on language structures. The indicator of gender intensity of a language that has the best coverage is one that relies on "gender distinctions in independent personal pronouns." Following prior work (Santacreu-Vasut et al. 2013), we use this indicator from WALS and code it into a binary variable, which equals "0" for the gendered languages, i.e., those that make gender distinctions in third-person pronouns and in the first and /or the second person. All other languages are coded as *Genderless languages* and keyed as "1" otherwise.[21]

Our analyses also include a number of covariates to assess the robustness of any language influence we uncover. We include a country-level control for a nations' level of democracy (*Democracy*), as well as individual-level controls for education (*Education*), income (*Income*), unemployment status (*Unemployed*), marital status (*Married*), gender (*Sex*), and age (*Age*).[22] All of our models also include fixed effects for survey waves. Furthermore, in at least one third of the countries included in our analysis, some respondents speak at home a gendered tongue, while others speak a genderless language. Hence, we also include country-fixed effects and use within-country variance to better identify language's influence. All models use robust standard errors clustered on country.

---

[21] This coding scheme contrasts strongly gendered languages with the weakly gendered and genderless ones. Our results also hold when we use a more nuanced coding with three categories of languages: genderless (2), weakly gendered (1), and strongly gendered (0). The details on this alternative measure and the corresponding robustness tests are presented in OA 7.

[22] The details on these measures are presented in OA.7.

Results from Study 3

The results from this cross-national analysis are presented in Table 3. The first model under each

dependent variable reports our most basic finding: the coefficient on *Genderless language*

captures the average difference in people's propensity to express gender-equal attitudes between

speakers of gendered and genderless languages. The coefficient is correctly signed and

statistically significant, such that individuals who speak a genderless language are more likely to

express gender-equal attitudes. Note that these models include country fixed effects. Thus, even

individuals living in the same country express reliably different attitudes about gender equality

depending on whether they speak a gendered or genderless language at home. This result closely

mirrors our main experimental findings.

This basic pattern is robust to individual- and country-level controls (the second model

under each dependent variable). In the OLS models, the coefficients run from a low of 0.12 to a

high of 0.32, which imply shifts of about 3% to 8% in those dependent variables. In the logit

model (DV=*Women jobs*), the implied language effect is substantively larger. There, a

respondent who reports speaking a genderless language at home is about 60% more likely to

disagree with this statement than a respondent who reports speaking a gendered language. These

type of findings are consistent with cross-national non-survey analyses suggesting that countries

that use genderless tongues experience greater levels of gender disparities in society (Prewitt-

Freilino et al. 2009).

In sum, our cross-national analysis corroborates our experimental results. We interpret

this as remarkable evidence of the strength and robustness of our argument about the effect of

language on attitudes toward gender equality. Even though an observational analysis of survey

data is not equivalent to a tightly controlled experiment, the fact that we were able to replicate

our basic finding in different research contexts, with different kinds of individuals and dependent variables, and a different operationalization of relevant language nuances, bolsters the external validity and generalizability of our experimental results.

Summary and Implications

We have argued that speaking a genderless tongue leads people to de-emphasize gender distinctions in one's perceptual ken. Because of this, speakers of genderless tongues find it harder to sense a "natural" asymmetry between men and women, which leads them to be more supportive of efforts to address gender inequality. We tested this proposition in three independent studies. In the first one, we manipulated whether bilinguals interviewed in a genderless or gendered tongue, thus allowing us to isolate our proposed mechanism (grammatical gender). Using this design, we find that those assigned to interview in a genderless language are significantly less likely to view women in a stereotypical way and significantly more likely to support efforts to increase gender equality. The boost in the probability of support triggered by interviewing in a genderless language ranges between five to eight percentage points. These results do not appear to be chance occurrences, for we successfully replicated them in Study 2, which used the same type of experimental design in a fresh sample of bilingual respondents.

We also established an important boundary condition to our hypothesized effect: speaking a genderless tongue has a trivial influence on one's opinions about gender inequality when the domain in question is attended by clearer social norms about appropriate behavior. Such norms provide additional information that plug mental gaps that would otherwise be filled by one's tongue. We further established that the effects of speaking genderless language fail to emerge in domains that do not evoke gender distinctions – a result that further validates our

causal claim. Finally, we demonstrated the external validity of our language effects through an extensive cross-national analysis of survey data, finding that speakers of genderless languages hold more liberal political attitudes toward gender equality than speakers of gendered languages. Combined, these experimental and observational results paint a clear picture of language's political relevance.

By highlighting language differences in support for efforts to reduce gender inequality, we also help to illuminate its persistence. Our findings reveal that some people cannot be more supportive of efforts to more fully reach this goal simply because of the language they speak. In this way, we have shown that language can operate as a ceiling to some of the beneficial trends for gender equality unleashed by higher socio-economic development (e.g., flexible family leave policy).

Of course, if something so fundamental as the language spoken by people can color their views of gender equality, then what is there to do—besides throwing up one's arms—about the kinds of differences we have detected between gendered and non-gendered languages? We believe that the boundary condition we established offers a potential solution. Recall that language has a negligible impact on people's views about gender equality when norms about appropriate behavior are clear and strong. This insight opens the door to social pressure, more generally, as a possible moderator of the types of language effects we have reported. Furthermore, our results lend at least some credence to the efforts of adopting gender-neutral terms in different languages. Such efforts, our result suggest, amount to more than just window-dressing or "political correctness." Rather, they can have real consequences for how salient gender categories are to speakers and, as a result, to their attitudes about gender equality.

In the end, we have peeled back a layer of people's minds to show how the language one speaks can influence what one sees, and fails to see, when it comes to equality between men and women. Language, we have learned, is a key contributor to why we might still see gender inequality where more progress is expected. As such, our theoretical framework is psychological, through and through. It therefore only illuminates the deep corners of people's minds. But those recesses are important, our results suggest, because simple quirks in language can affect how mentally receptive people can be to an even more equitable balance between men and women.

References

Adida, Claire L., Karen E. Ferree, Daniel N. Posner, and Amanda L. Robinson. 2016. "Who's Asking? Interviewer Coethnicity Effects in African Survey Data." *Comparative Political Studies* forthcoming.

Almond, Gabriel A., and Sidney Verba. 1963. *The Civic Culture Political Attitudes and Democracy in Five Nations*. Princeton: Princeton University Press.

Arriola, Leonardo R., and Martha C. Johnson. 2014. "Ethnic Politics and Women's Empowerment in Africa: Ministerial Appointments to Executive Cabinets." *American Journal of Political Science* 58 (2): 495-510.

Blau, Francine D., Mary C. Brinton, and David B. Grusky. 2006. *The Declining Significance of Gender?* New York: Russell Sage Foundation.

Boroditsky, Lera. 2000. "Metaphoric Structuring: Understanding Time Through Spatial Metaphors." *Cognition* 75: 1-28.

Boroditsky, Lera. 2001. "Does Language Shape Thought?: Mandarin and English Speakers' Conception of Time." *Cognitive Psychology* 43(1): 1-22.

Boroditsky, Lera and Alice Gaby. 2010. "Remembrances of Times East: Absolute Spatial Representations of Time in an Australian Aboriginal Community." *Psychological Science* 21(11): 1635-1639.

Boroditsky, Lera, Lauren A. Schmidt, and Webb Phillips. 2003. "Sex, Syntax, and Semantics." In *Language in Mind: Advances in the Study of Language and Cognition*, ed. Dedre Gentner and Susan Goldin-Meadow. Cambridge: MIT Press.

Bueno de Mesquita, Bruce, and Alastair Smith. 2012. "Domestic Explanations of International Relations." *Annual Review of Political Science* 15: 161-181.

Campbell, Donald T., and Julian C. Stanley. 1963. *Experimental and Quasi-Experimental Designs for Research*. Boston: Houghton Mifflin Company.

Corbett, Greville G. 1991. *Gender*. Cambridge: Cambridge University Press.

Cubelli, Roberto, Daniela Paolieri, Lorella Lotto, and Remo Job. 2011. "The Effect of Grammatical Gender on Object Categorization." *Journal of Experimental Psychology* 37(2): 449-460.

Dalton, Russell J. 2014. *Citizen Politics*. Washington: CQ Press.

Danziger, Shai and Robert Ward. 2010. "Language Changes Implicit Associations Between Ethnic Groups and Evaluation in Bilinguals." *Psychological Science* 21(6): 799-800.

Dryer, Matthew S. and Martin Haspelmath (eds.). 2013. *The World Atlas of Language Structures Online.* Leipzig: Max Planck Institute for Evolutionary Anthropology.

Dunning, Thad. 2016. "What Experiments Can and Cannot Achieve in Political Science." *Annual Review of Political Science* 19: 541-563.

Echterhoff, Gerald. 2008. Language and Memory: Social and Cognitive Processes. In A. Erll and A. Nünning, eds., *Cultural Memory Studies: An International and Interdisciplinary Handbook*. Berlin: Walter de Gruyter GmbH & Company.

Epstein, Cynthia Fuchs. 2007. "Great Divides: The Cultural, Cognitive, and Social Bases of the Global Subodination of Women." *American Sociological Review* 72(1): 1-22.

Fausey, Caitlin, Bria L. Long, Aya Inamori, and Lera Boroditsky. 2010. "Constructing Agency: The Role of Language." *Frontiers in Psychology* 1(162): 1-11.

Fortin, Nicole M. 2005. "Gender Role Attitudes and the Labour-market Outcomes of Women across OECD Countries." *Oxford Review of Economic Policy* 21(3): 416-438.

Fuhrman, Orly, Kelly McCormich, Eva Chen, Heidi Jiang, Dingfang Shu, Shuaimei Mao, and Lera Boroditsky. 2011. "How Linguistic and Cultural Forces Shape Conceptions of Time: English and Mandarin Time in 3D." *Cognitive Science* 35: 1305-1328.

Garcia, John A. 2009. "Language of Interview: Exploring Patterns and Consequences of Changing Language During Interview." Paper presented at the annual meeting of the Western Political Science Association. Vancouver, B.C.

Garcia Bedolla, Lisa. 2005. *Fluid Borders: Latino Power, Identity, and Politics in Los Angeles*. Berkeley: University of California Press.

Green, Donald P. 2004. "Experimental Design." In M.S. Lewis-Beck, A. Bryman, and T.F. Liao, eds.,*The Sage Encyclopedia of Social Science Research Methods*. Thousand Oaks: Sage Publications.

Guiora, Alexander Z., Benjamin Beit-Hallahmi, Risto Fried, and Cecelia Yoder. 1982. "Language Environment and Gender Identity Attainment." *Language Learning* 32(2): 289-304.

Hinojosa, Magda. 2012. *Selecting Women, Electing Women: Political Representation and Candidate Selection in Latin America.* Philadelphia: Temple University Press.

Hicks, Daniel L., Estefania Santacreu-Vasut, and Amir Shoham. 2015. "Does Mother Tongue Make for Women's Work? Linguistics, Household Labor, and Gender Identity." *Journal of Economic Behavior and Organization* 110: 19-44.

Hopkins, Daniel. 2014. "One Language, Two Meanings: Partisanship and Responses to Spanish." *Political Communication* 31(3): 421-445.

Hopkins, Daniel. 2015. "The Upside of Accents: Language, Inter-group Difference, and Attitudes toward Immigration." *British Journal of Political Science* 45(3): 531-557.

Hopkins, Daniel, Van C. Tran, and Abigail Fisher Williamson. 2014. "See no Spanish: Language, Local Context, and Attitudes Toward Immigration." *Politics, Groups, and Identities* 2(1): 35-51.

Horowitz, Donald L. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.

Huddy, Leonie and Nayda Terkildsen. 1993. "Gender Stereotypes and the Perception of Male and Female Candidates." *American Journal of Political Science* 37(1): 119-147.

Hunt, E. and F. Agnoli. 1991. "The Worfian Hypothesis: A Cognitive Psychology Perspective." *Psychological Review* 98(3): 377-389.

Inglehart, Ronald and Pippa Norris. 2003. *Rising Tide: Gender Equality and Cultural Change Around the World*. Cambridge: Cambridge University Press.

Iversen, Torben and Frances Rosenbluth. 2010. *Women, Work, and Politics: The Political Economy of Gender Inequality*. New Haven: Yale University Press.

Kay, Paul and Willett Kempton. 1984. "What Is the Sapir-Whorf Hypothesis?" *American Anthropologist* 86(1): 65-79.

Kinder, Donald R. and Cindy D. Kam. 2009. *Us Against Them: Ethnocentric Foundations of American Opinion*. Chicago: University of Chicago Press.

Koch, Jeffrey. 2002. "Gender Stereotypes and Citizens' Impression of House Candidates Ideological Orientations." *American Journal of Political Science* 46: 453-462.

Konishi, Toshi. 1993. "The Semantics of Grammatical Gender: A Cross-Cultural Study." *Journal of Psycholinguistic Research* 22(5): 519-534.

Laitin, David D. 1998. *Identity in Formation.* Ithaca: Cornell University Press.

Laitin, David D. 1977. *Politics, Language, and Thought: The Somali Experience*. Chicago: University of Chicago Press.

Laitin, David D., Joachim Moortgat, and Amanda Lea Robinson. 2012. "Geographic Axes and the Persistence of Cultural Diversity." *PNAS* 109(26): 10263-10268.

Lakoff, George. 1987. *Women, Fire, and Dangerous Things.* Chicago: University of Chicago Press.

Lauristin, Marju, and Triin Vihalemm. 1997. "Changing Value Systems: Civilizational Shift and Local Differences." In *Return to the Western World: Cultural and Political Perspectives on the Estonian Post-Communist Transition*, edited by M. Lauristin and P. Vihalemm. Tartu: Tartu University Press.

Lodge, Milton and Charles S. Taber. 2013. *The Rationalizing Voter*. New York: Cambridge University Press.

Maimone, Christina. 2004. "The Estonian Russian Divide: Examining Social Diversity in Estonia with Cross-National Survey Data." Working Paper, Stanford University, http://socsci2.ucsd.edu/~aronatas/scrretreat/Maimone.Christina.pdf

Marian, Viorica and Ulirc Neisser. 2000. "Language-Dependent Recall of Autobiographical Memories." *Journal of Experimental Psychology: General* 129(3): 361-368.

O'Brien, Diana Z. 2015. "Rising to the Top: Gender, Political Performance, and Party Leadership in Advanced Industrial Democracies." *American Journal of Political Science* 59(4): 1022-1039.

O'Brien, Diana Z. and Johanna Rickne. 2016. "Gender Quotas and Women's Political Leadership." *American Political Science Review* 110(1): 112-126.

OECD. 2012. Gender Equality in Education, Employment and Entrepreneurship: Final Report to the MCM 2012.

Paxton, Pamela, Melanie M. Huges, and Matthew A. Painter II. 2010. "Growth in Women's Political Representation: A Longitudinal Exploration of Democracy, Electoral System and Gender Quotas." *European Journal of Political Research* 49(1): 25-52.

Phillips, Webb and Lera Boroditsky. 2003. "Can Quirks of Grammar Affect the Way You Think? Grammatical Gender and Object Concepts." In Proceedings of the 25th Annual Meeting of the Cognitive Science Society, ed. Richard Alterman and David Kirsh. Boston, MA: Cognitive Science Society.

Posner, Daniel N. 2005. *Institutions and Ethnic Politics in Africa.* New York: Cambridge University Press.

Prewitt-Freilino, Jennifer L., T. Andrew Caswell, 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(3): 268-281.

Sanbonmatsu, Kira. 2002. "Gender Stereotypes and Vote Choice." *American Journal of Political Science* 46(1): 20-34.

Santacreu-Vasut, Estefania, Amir Shoham, and Victor Gay. 2013. "Do Female/Male Distinctions in Language Matter? Evidence from Gender Political Quotas. *Applied Economic Letters* 20: 495-498.

Senden, Marie Gustafsson, Emma A Bäck, and Anna Lundqvist. 2015. "Introducing a Gender-neutral Pronoun in a Natural Gender Language: The Influence of Time on Attitudes and Behavior." *Frontiers in Psychology* 6:893. doi: 10.3389/fpsyg.2015.00893

Sera, Maria D., Christian A.H. Berge, and Javier del Castillo Pintado. 1994. "Grammatical and Conceptual Forces in the Attribution of Gender by English and Spanish Speakers." *Cognitive Development* 9(3): 261-292.

Slobin, Dan. 2003. "Language and Thought Online: Cognitive Consequences of Linguistic Relativity." In D. Gentner and S. Goldin-Meadow, eds., *Language in Mind: Advances in the Study of Language and Thought*. Boston: Massachusetts Institute of Technology Press.

Slobin, Dan. 1996. "From "Thought and Language" to "Thinking for Speaking"." In *Rethinking Linguistic Relativity*, ed. John J. Gumperz and Stephen C. Levinson. Cambridge: Cambridge University Press.

Sobolewska, Maria, Laurence Lessard-Phillips, and Silvia Galandini. 2016. "Exploring Public Opinion on Immigrant Integration with Survey Experiments." *Migration and Citizenship* 4(1): 18-23.

Taber, Charles S., and Everett Young. 2013. "Political Information Process." In L. Huddy, D.O. Sears, and J.S. Levy, eds., *The Oxford Handbook of Political Psychology*. New York: Oxford University Press.

Tourangeau, Roger, Lance J. Rips, Kenneth Rasinski. 2000. *The Psychology of Survey Response*. New York: Cambridge University Press.

Valentino, Nicholas A. and Yioros Nardis. 2013. "Political Communication: Form and Consequence of the Information Environment." In Leonie Huddy, David O. Sears, and Jack S. Levy (eds.) *Oxford Handbook of Political Psychology*. Oxford University Press, 559-590.

Vigliocco, Gabriella, David P. Vinson, Federica Paganelli, Katharina Dworzynski. 2005. "Grammatical Gender Effects on Cognition: Implications for Language Learning and Language Use." *Journal of Experimental Psychology: General* 134(4): 501–520.

Winawer Jontahan, Nathan Witthoft, Michael C. Frank, Lisa Wu, Alex R. Wade, and Lera Boroditsky. 2007. "Russian Blues Reveal Effects of Language on Color Discrimination." *Proceedings of National Academy of Sciences* 104: 7780-7785.

Zaller, John. 1992. *The Nature and Origins of Mass Opinion.* New York: Cambridge University
Press.

Table 1. The Effect of Genderless Language on Opinions Toward Gender Equality (Study 1)

	Model 1: Emotional women: relative rating (OLS)	Model 2: Paternity Leave (Probit)	Model 3: Female Defense Minister (Probit)	Model 4: Female political recruitment (Ordered Probit)
Estonian interview	-0.20*	0.21**	0.22**	0.14**
	(0.12)	(0.08)	(0.08)	(0.06)
Constant	1.34***	-0.38***	0.40***	---
	(0.09)	(0.05)	(0.05)	
N	1,153	1,140	1,156	1,154

Note: Dependent variables are indicated in column headings. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$, two-tailed tests.

Table 2. The Effect of Genderless Language on Opinions Toward Gender Equality (Study 2)

	Replication			Social Norms Experiment		Placebo Test
	Model 1: Paternity Leave (Probit)	**Model 2: Female Defense Minister** (Probit)	**Model 3: Female political recruitment** (Ordered probit)	**Model 5: Run for office– no norm** (Ordered probit)	**Model 6: Run for office– with norm** (Ordered probit)	**Model 7: Suicide placebo item** (OLS)
Estonian interview	0.41** (0.16)	0.26* (0.19)	0.01 (0.14)	0.36** (0.21)	0.22 (0.22)	0.15 (0.28)
Constant	-0.20 (0.11)	0.74** (0.12)	---	---	---	2.14** (0.19)
N	251	248	253	134	123	243

Note: Dependent variables are indicated in column headings. **$p<0.05$, *$p<0.10$, one-tailed tests.

Table 3. The Effect of Genderless Language on Opinions Toward Gender Equality, World Values Survey 1995-2014 (Study 3)

	Women Polit. Leader		Univ. for Girls		Women Busin. Exec.		Women Jobs		Gender Equal Attitudes	
	Model 1 (OLS)	Model 2 (OLS)	Model 1 (OLS)	Model 2 (OLS)	Model 1 (OLS)	Model 2 (OLS)	Model 1 (logit)	Model 2 (logit)	Model 1 (OLS)	Model 2 (OLS)
Genderless	0.204***	0.181***	0.324***	0.321***	0.149**	0.121**	0.611***	0.471***	0.254***	0.216**
language	(0.046)	(0.067)	(0.111)	(0.100)	(0.057)	(0.057)	(0.071)	(0.099)	(0.091)	(0.090)
Sex		0.281***		0.232***		0.323***		0.551***		0.357***
		(0.014)		(0.018)		(0.022)		(0.041)		(0.024)
Age		-0.002***		-0.002***		-0.002***		0.007***		-0.002***
		(0.000)		(0.001)		(0.000)		(0.002)		(0.001)
Unemployed		0.025**		-0.027***		0.014		0.075**		-0.001
		(0.011)		(0.010)		(0.014)		(0.035)		(0.015)
Married		-0.025***		-0.008		-0.014		-0.152***		-0.030**
		(0.008)		(0.008)		(0.010)		(0.031)		(0.012)
Democracy		0.002		0.009		0.005		0.010		-0.017
		(0.009)		(0.010)		(0.017)		(0.014)		(0.020)
Country FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Wave FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Education		YES		YES		YES		YES		YES
Income		YES		YES		YES		YES		YES
Constant	2.173***	2.252***	2.656***	2.644***	2.506***	2.584***	0.160	0.178***	-0.192***	-0.004
	(0.041)	(0.082)	(0.067)	(0.080)	(0.037)	(0.099)	(0.106)	(0.222)	(0.058)	(0.117)
N	168,903	135,827	172,828	138,761	100,373	83,324	176,939	141,632	95,862	79,937
N (countries)	90	84	90	84	75	71	90	83	75	71
R-squared	0.004	0.041	0.005	0.040	0.001	0.049			0.002	0.072

Note: Dependent variables are indicated in column headings. Robust country-clustered standard errors in parentheses. ***$p<0.01$, **$p<0.05$, *$p<0.1$

Figure 1. Probability of Support for Political Efforts Addressing Gender Inequality

A. Change in Paternity Leave Policy

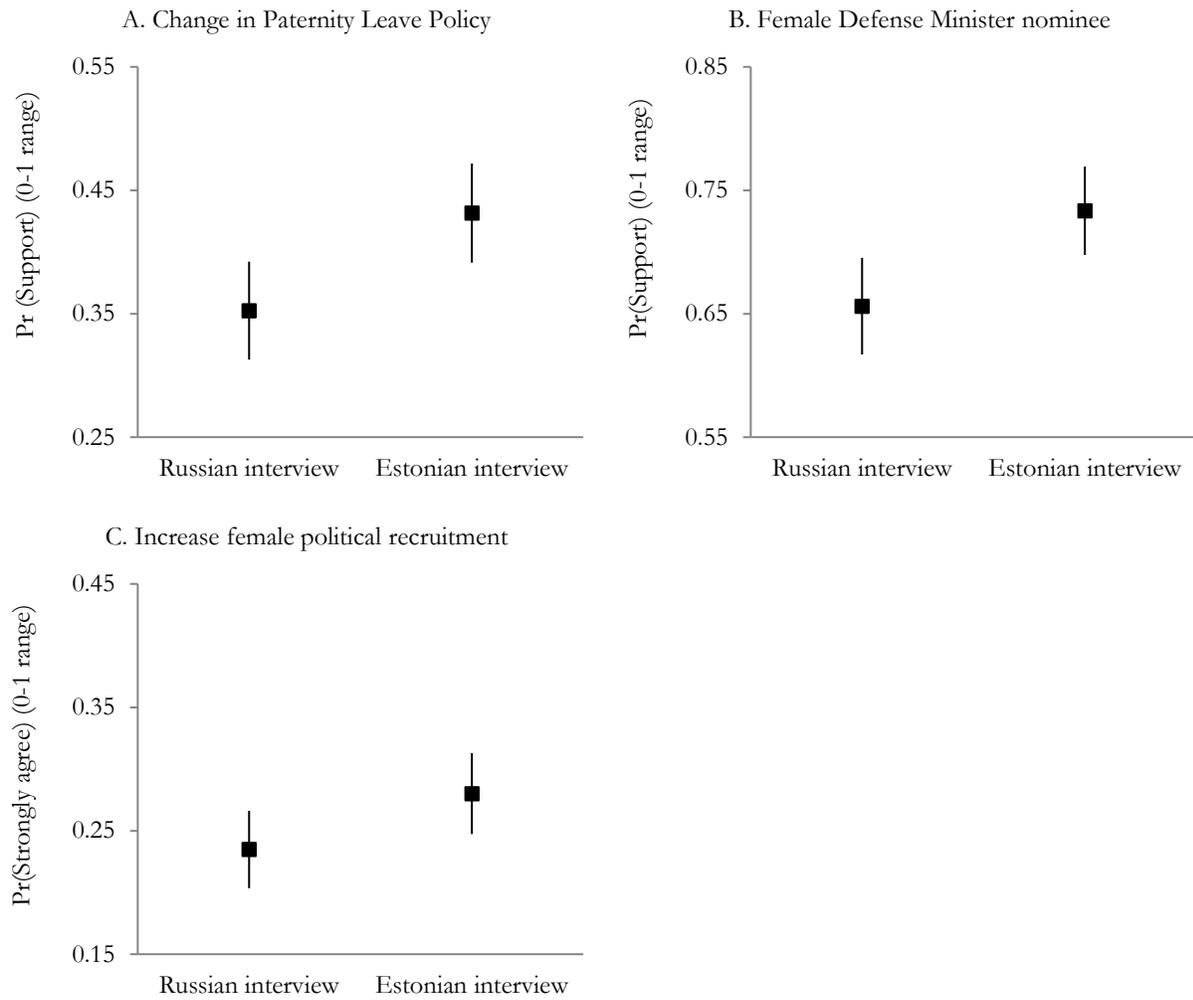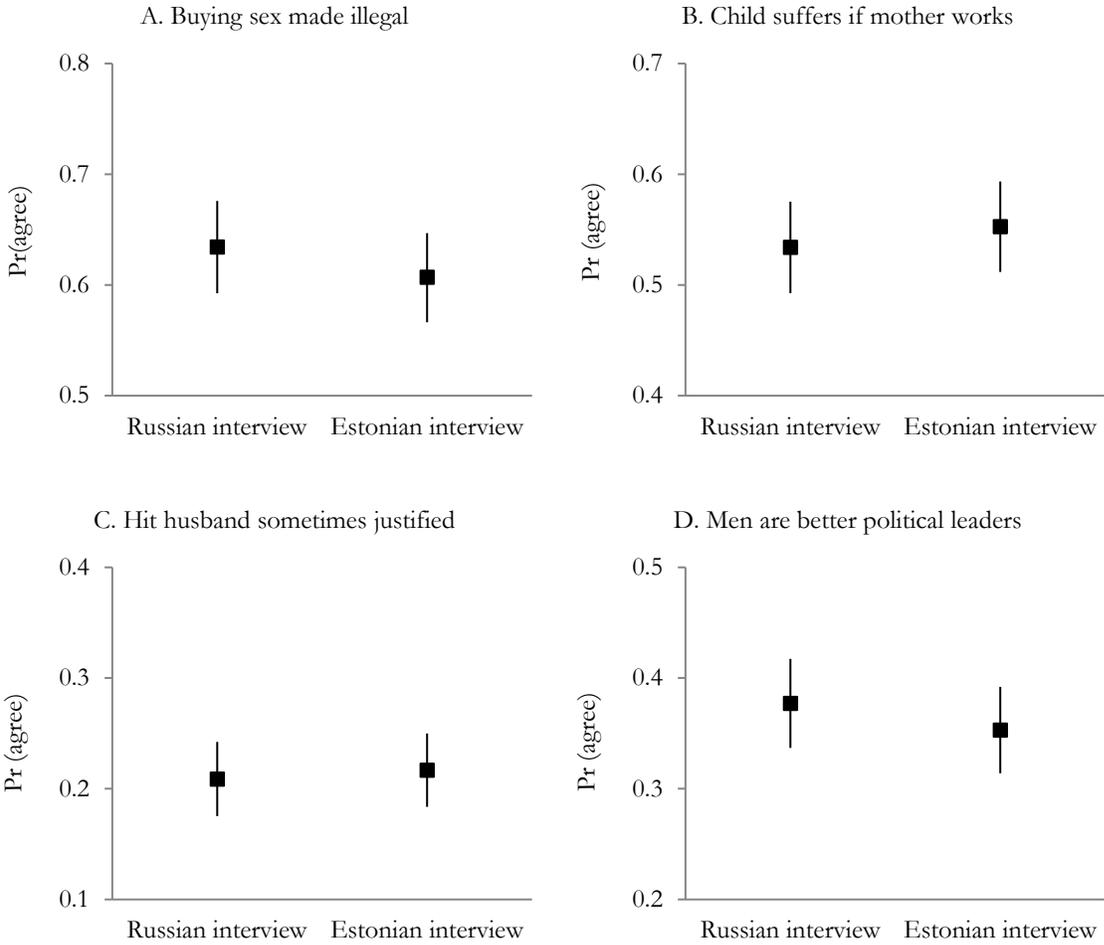B. Female Defense Minister nominee

C. Increase female political recruitment

Figure 2. Boundary Condition: Language Effects Are Less Likely When Social Norms Clearly Imply Appropriate Behavior

A. Buying sex made illegal

B. Child suffers if mother works

C. Hit husband sometimes justified

D. Men are better political leaders

Online Appendix

"Language Shapes Public Attitudes Toward Gender Equality"

This Online Appendix file contains additional information on the survey experiments, i.e., Studies 1 and 2 (OA.1), randomization and balance checks for Study 1 (OA.2), robustness tests for Study 1 (OA.3), the analysis probing into whether the language effects reveal pro-female bias or gender-blindness (OA.4), raw regression results for the analysis of boundary conditions (Study 1) (OA.5), randomization and balance checks for Study 2 (OA.6), and additional information on the cross-national analysis (Study 3) (OA.7).

OA.1. Additional Information about Studies 1 and 2

Studies 1 and 2 were administered via telephone by TNS Emor, a leading Estonian survey firm with extensive experience conducting survey research for public and private sector clients, including the Eurobarometer Survey (since 2004). Our study's universe includes Estonian residents, ages 18-74, who can speak Estonian and Russian. The sampling frame consisted of randomly drawn landline and mobile phone numbers.

OA.1.1 Language Manipulations
In both Study 1 and Study 2, our language manipulation was designed to set a distinct linguistic milieu by informing bilinguals that all subsequent instructions and questions would be in their assigned tongue.[1] This effort was further bolstered by the fact that our bilingual interviewers were trained to strictly engage respondents in the assigned interview language. Specifically, our language treatment in Study 1 was worded as follows:

> "Based on your answers to some of the previous questions, it appears that you are fluent in both Estonian and Russian. Therefore, we will let the computer program randomly select which language we continue this interview in [SHORT PAUSE].
>
> [*Estonian/Russian*] was selected. This means that after this point, the rest of the interview will take place in [*Estonian/Russian*]. This is not a language test. We are simply interested in your opinions as an [*Estonian/Russian*] speaker."

In Study 2, the treatment followed the same interviewer script as above, except it nixed the last line, which stated: "We are simply interested in your opinions as an [*Estonian/Russian*] speaker." We made this change to eliminate any lingering concern that what we manipulated in Study 1 was language identity, rather than language of interview.

OA.1.2 Translation and Interviewer Protocols
Our treatments and survey items I both studies were designed in (or adapted from) English by both authors, and then translated into Estonian and Russian by TNS Emor in collaboration with one author. Before fielding Study 1, we conducted an independent pre-test of the translated items, which failed to find any further language issues (N = 88). We note, however, that we did not

---

[1] Marian, Viorica and Ulirc Neisser. 2000. "Language-Dependent Recall of Autobiographical Memories." *Journal of Experimental Psychology: General* 129(3): 361-368.

seek to create translations that were exact in length, word for word. Such a strategy can yield translations of different and grammatically incorrect meanings, thus rendering any language comparisons useless.[2] Instead, we prioritized the development of translated questions that meant the same thing to different language speakers, since detecting real opinion differences assumes that people share the same notion of what is being asked.

In the case of both Study 1 and 2, live bilingual interviewers conducted our phone surveys via CATI (i.e., computer-assisted telephone interviewing). Although TNS Emor employs male interviewers, the number of *bilingual* male interviewers at the time of our studies was too low ($n = 5$) for us to either (a) efficiently block our treatment on interviewer gender; or (b) statistically adjust our treatment effects for observed interviewer gender. Thus, we exclusively relied on female bilingual interviewers whose first language was Estonian ($n = 41$), which matches the language profile of most of our bilingual respondents. This choice means our respondents might feel obliged to give female interviewers pro-woman responses. But if social desirability bias is really present, randomization ensures that it will be equal across our language conditions. Moreover, if it has any effect, such pressure is likely to work against finding opinion differences between our interview groups. That is, because all respondents may feel obliged to give pro-women responses, any opinion gap between interviewees will be smaller than what would emerge in the absence of social desirability, thereby making our estimates conservative ones. Another possibility is that our interviewers lead male (female) respondents to express stronger anti-female opinions because they feel that women "should not" be working in visibly public roles, like interviewing survey respondents. If this is true, then respondents' gender should consistently moderate the effect of interview language on opinions. It does not (see Table OA.3.1).

OA.1.2 Question Wording, Study 1

Emotional women/Emotional men
[ITEMS 'A' AND 'B' IN RANDOMIZED ORDER]

a. Where would you rate men in general on a scale of 1 to 7? (where 1 indicates rational, 7 means emotional, and 4 indicates most men are not closer to one end or the other.

b. Where would you rate women in general on a scale of 1 to 7? (where 1 indicates rational, 7 means emotional, and 4 indicates most men are not closer to one end or the other.

Paternity leave and Female Defense Minister
[ITEMS 'C' AND 'D' IN RANDOMIZED ORDER]

c. Under the current legislation, the father can take paid parental leave only after the baby is 2 months old. Do you agree or disagree that this policy should be changed so that the father could stay home with the baby and the mother could return to work from the day the baby is born.
    1. Agree
    2. Disagree

---

[2] Jakobson, Roman. 1966. "On Linguistic Aspects of Translation." In R. A. Brower (Ed.) *On Translation.* New York: Oxford University Press. Pp. 232-239.

d. If the party that you normally like nominated a generally well-qualified woman to be Minister of Defense, would you support that choice?
  1. Yes, I would support that choice.
  2. No, I would oppose that choice.

Female political recruitment and *Female candidates: men's expense*
Next, I am going to read you some proposed government efforts to address several social issues. After I read each one, please tell me whether you strongly disagree, somewhat disagree, somewhat agree, or strongly agree with each statement. What about…

[ITEMS 'E' AND 'F' IN RANDOMIZED ORDER]

e. Require political parties to reserve some space on their lists of candidates for women, even if they have to exclude some men.
  1. Strongly disagree
  2. Somewhat disagree
  3. Somewhat agree
  4. Strongly agree

f. And, what about: Recruit more women to top-level government positions.
  1. Strongly disagree
  2. Somewhat disagree
  3. Somewhat agree
  4. Strongly agree

Buy sex, Child suffers, Hit husband, Men better
g. Under the current law, "pimping" is a criminal act but buying sex is not. Do you agree or disagree with the statement: Buying sex should also be made illegal.
  1. Agree
  2. Disagree

h. What about the statement: A preschool child is likely to suffer if his or her mother works.
  1. Agree
  2. Disagree

i. And what about the statement: It is sometimes justified for a woman to hit her husband.
  1. Agree
  2. Disagree

j. And what about the statement: On the whole, men make better political leaders than women do.
  1. Agree
  2. Disagree

OA.1.3 Question Wording, Study 2

The wording of *Paternity leave*, *Female Defense Minister*, and *Female political recruitment* is the same as in Study 1. Here is the wording of the two additional questions that we included in Study 2.

Run for office
iv. Next, I am going to read you some proposed government efforts to address several social issues. After I read each one, please tell me whether you strongly disagree, somewhat disagree, somewhat agree, or strongly agree with each statement. What about…

[RANDOMLY ASSIGN ONE ITEM 1a or 1b]

1a. Calling on party leaders to encourage more women to run for office.

[OR]

1b. Calling on party leaders to encourage more women to run for office, a proposal that about 80% of the people in Estonia favor.

    1. Strongly disagree
    2. Somewhat disagree
    3. Somewhat agree
    4. Strongly agree

Suicide
Next, on a scale of 1 to 10, please tell me the degree to which you think suicide is justifiable, where "1" is never justifiable and "10" is always justifiable.

    1) Never justifiable
    2)
    3)
    4)
    5)
    6)
    7)
    8)
    9)
    10) Always justifiable.

OA.2. Randomization and Balance Checks, Study 1

We measured several pre-treatment covariates: respondents' age, education level, gender, ideology, first language learned, and preferred language of interview. Table OA.2.1 uses these covariates to predict our respondent's assignment to interview in Estonian or Russian. A Wald test shows we cannot reject the null that these covariates are simultaneously equal to zero ($\chi^2 = 8.50$, Prob χ^2 (8) $> 8.50 = 0.39$), which is consistent with the random assignment of interview language.

Table OA.2.1. Randomization Check: Language Manipulation is Unrelated to Pre-Treatment Covariates

	Estonian Interview (=1)
Education	0.006
	(0.037)
Female	-0.115
	(0.076)
Age	-0.005
	(0.003)
Left	-0.145
	(0.126)
Right	-0.095
	(0.100)
Center	-0.087
	(0.092)
Russian first	0.049
	(0.121)
Prefer Russian	-0.082
	(0.124)
Constant	0.397
	(0.214)
N	1,200

Note: Entries are probit coefficients with standard errors in parentheses. An omnibus Wald test shows that one cannot reject the null that the covariates in Table OA.2.1 are simultaneously equal to zero, i.e., unassociated with the language manipulation ($\chi^2 = 8.50$, Prob χ^2 (8) $> 8.50 = 0.39$).

The range and units of the covariates are as follows:
- *Education* runs from 1-Elementary to 5-University in single units.
- *Female* is a dummy variable where males are the omitted category.
- *Age* ranges from 18 to 74 in years.
- *Left*, *Right*, and *Center* are dummy variables with "don't knows" and "refusals" as the omitted category. *Left* indicates R placed themselves on the left side of a 10-point left-right ideology scale, while *Right* indicates R placed themselves on the right side of that same scale. *Center* indicates R placed themselves at the midpoint of this ideology scale.
- *Russian first* indicates bilingual R learned Russian first.
- *Prefer Russian* indicates bilingual R prefers to interview in Russian.

We also examined whether these covariate distributions are similar across our experimental groups (i.e., balance check). As with any experiment, *chance* variations in these distributions are anticipated: one or two reliable differences here should not be surprising, especially given our high-powered study. Yet we should not find large and reliable imbalances across most of these covariates. Table OA.2.2 reveals that seven (7) out of eight (8) covariate distributions are statistically identical across our experimental groups. Only one reliable difference emerges ($p < 0.05$, two-tailed), where the distribution of Russian interviewees is a bit older (median = 53 years) than Estonian interviewees (median = 50). Given our large sample, we interpret this gap as statistically significant but substantively negligible. There is also a small difference between the female proportion in the Estonian (59%) and Russian (63%) conditions that approaches marginal significance ($p < 0.103$). But the slightly greater female percentage emerges among Russian interviewees, who report their opinions in a gendered tongue. This implies that, if anything, it will be even harder to detect opinion differences between Estonian and Russian interviewees.

Table OA.2.2. Distributions of Pre-Treatment Covariates (Balance Check)

	Estonian interview	Russian interview	Chi-square test	Wilcoxon-Mann-Whitney test
Education (median)	3	3	---	$z = 0.029$ Prob $>|z| = 0.977$
Female (%)	59%	63%	$\chi^2(1) = 2.66$, $p < 0.103$	---
Age (median)	50	53	---	$z = 1.981$ Prob $> |z| = 0.048$
Left (%)	10%	11%	$\chi^2(1) = 0.641$ $p < 0.423$	---
Right (%)	24%	24%	$\chi^2(1) = 0.065$ $p < 0.799$	---
Center (%)	26%	27%	$\chi^2(1) = 0.176$ $p < 0.675$	---
Russian first (%)	38%	38%	$\chi^2(1) = 0.083$ $p < 0.772$	---
Prefer Russian (%)	31%	31%	$\chi^2(1) = 0.024$ $p < 0.876$	---

Note: The differences in statistical tests performed reflect the nature of the covariates under analysis (e.g., dichotomous versus interval). All significance tests are two-tailed.

OA.3. Robustness Tests, Study 1

Table OA.3.1. Estimated Treatment Effects With Covariates Included

A.	Emotional Women (relative)	Emotional Women (single)	Emotional Men (single)	Paternity Leave	Female Defense Minister	Female Political Recruit.	Females at Men's Expense
Estonian interview	-0.19[†]	-0.22**	-0.02	0.18**	0.20**	0.15**	-0.14**
	(0.11)	(0.09)	(0.08)	(0.08)	(0.08)	(0.06)	(0.06)
Age	-0.45*	-0.50**	-0.03	-0.82**	-0.76**	0.07	0.00
	(0.24)	(0.18)	(0.17)	(0.16)	(0.16)	(0.12)	(0.00)
Education	-0.69**	0.11	-0.58**	-0.32**	0.48**	-0.47**	-0.15**
	(0.24)	(0.17)	(0.16)	(0.15)	(0.16)	(0.13)	(0.03)
Female	0.44**	0.13	-0.29**	-0.04	-0.06	0.24**	0.18**
	(0.12)	(0.09)	(0.08)	(0.08)	(0.08)	(0.07)	(0.07)
Constant	0.85**	4.94**	4.07**	0.35**	0.58**	---	---
	(0.25)	(0.18)	(0.17)	(0.16)	(0.16)		
N	1,153	1,165	1,158	1,140	1,156	1,154	1,143

B.	Buying Sex	Child Suffers	Hit Husband	Men Better Leaders
Estonian interview	-0.05	0.07	0.01	-0.05
	(0.08)	(0.08)	(0.08)	(0.08)
Age	0.33**	0.50**	-0.09	0.50**
	(0.16)	(0.15)	(0.17)	(.016)
Education	-0.35**	-0.77**	-0.58**	0.17
	(0.16)	(0.14)	(0.17)	(0.15)
Female	0.55**	0.07	-0.26**	0.04
	(0.08)	(0.08)	(0.09)	(0.08)
Constant	0.06	0.28*	-0.20	-0.75**
	(0.16)	(0.16)	(0.17)	(0.16)
N				
	1,071	1,130	1,160	1,135

Note: All covariates run from 0 to 1. Female is a dummy variable. **$p < 0.05$; *$p < 0.10$; †$p < 0.11$, two-tailed.

Table OA.3.2. Estimated Language Effects Adjusted for Respondents' Preferred Interview Language

	Emotional women: Relative rating (OLS)	Paternity Leave (Probit)	Female Defense Minister (Probit)	Female Political Recruitment (Ordered Probit)	Female Candidates: Men's Expense (Ordered Probit)
Estonian Interview	-0.20* (0.12)	0.21** (0.08)	0.22** (0.08)	0.14** (0.06)	-0.15** (0.06)
Prefer to interview in Russian	0.27** (0.13)	0.21** (0.08)	-0.24** (0.08)	-0.002 (0.07)	0.22** (0.07)
Constant	1.26** (0.10)	-0.45** (0.06)	0.48** (0.06)	---	---
N	1,153	1,140	1,156	1,154	1,143

Note: Prefer to interview in Russian is a dichotomous variable, with 0 indicating a preference for interviewing in Estonian. **$p < 0.05$ *$p < 0.10$, two-tailed.

Table OA.3.3. The Effect of Interview Language is Not Moderated by Being Male

	Emotional women	Paternity Leave	Female Defense Minister	Female Political Recruitment	Females at Men's Expense	Buying Sex Illegal	Child Suffers	Hit Husband	Men Better Leaders
Estonian interview	-0.16 (0.16)	0.12 (0.10)	0.32* (0.10)	0.11 (.08)	-0.15* (0.08)	-0.05 (0.10)	0.07 (0.10)	-0.02 (0.11)	-0.08 (0.10)
Male	-0.41* (0.18)	-0.01 (0.11)	0.23* (0.11)	-0.29* (0.10)	-0.18† (0.10)	-0.54* (0.12)	-0.04 (0.11)	0.24* (0.12)	-0.10 (0.11)
Estonian x Male	-0.05 (0.24)	0.20 (0.16)	-0.28† (0.16)	0.12 (0.12)	0.02 (0.13)	-0.02 (0.16)	-0.04 (0.15)	0.08 (0.16)	0.06 (0.16)
Constant	1.50* (0.11)	-0.37* (0.07)	0.32* (0.07)	---	---	0.55* (.07)	0.10 (.07)	-0.91* (0.08)	-0.28* (0.07)
N	1,153	1,140	1,156	1,154	1,143	1,071	1,130	1,160	1,135

Note: Emotional women is estimated via OLS. Models where constant is absent are estimated via ordered probit. All other models are estimated via probit.
*$p < 0.05$, two-tailed; $^\dagger p<0.10$, two-tailed.

9

OA.4. Gender-blindness or Pro-female Bias?

Speaking a gender-less language appears to have a liberalizing effect on political attitudes toward gender equality, especially when social norms do not provide crisp and clear guidance on a matter. We claim that the mechanism responsible for these effects is the greater parity between gendered objects sensed by speakers of a gender-less tongue; let's call this mechanism "gender-blindness." But an alternative account is plausible: speakers of a gender-less language may be altogether more pro-female so that the improvement in gender parity is achieved at the expense of men; let's call this mechanism "pro-female bias." Clarifying which of these mechanisms is behind our observed effects is critical because it allows us to say whether greater gender balance is reached because the language makes speakers gender-blind or because it makes them see the female gender more positively than the male. The former, we believe, is a more accurate reflection of true gender equality, since female strides in society and politics are unrelated (or weakly related) to men's fortunes in these areas. The latter, in contrast, is a zero-sum bargain where female gains entail male displacement. To examine which of these mechanisms better explains the pattern of language effects we have uncovered, we return to two sets of dependent variables that we designed precisely to disentangle these alternatives.

The first set contains the rating of men and women as stereotypically emotional rather than rational. Recall that in their original format, respondents answered these items individually, sequentially, and—most importantly— in random order (variable names *Emotional women: single rating* and *Emotional men: single rating*). We have here, then, an opportunity to determine how respondents independently rated men and women, without the contamination of order effects. This allows us a strong test of the degree to which people's view of women is linked to their perception of men.

If the mechanism behind our observed language effects is "gender-blindness," then a more positive evaluation of females should be *unrelated* to how respondents interviewing in a gender-less language judge men. Empirically, this suggests a negative and reliable coefficient for the relationship between interviewing in Estonian and ratings of women "emotional," yet a small and statistically insignificant coefficient for ratings of men on this stereotype dimension. However, if "pro-female bias" is the correct mechanism then Estonian interviewees should judge women more positively on this stereotype dimension while rating men more negatively. This implies that Estonian interviewees are rating women as less stereotypically emotional while rating men as more stereotypically rational. Hence, we should observe a reliable negative and positive coefficient, respectively, which would reflect the relationship between interviewing in Estonian and stereotype judgments of females and males. Of course, one might argue that rating males as rational is not really a negative judgment; that it might, in fact, be more of a compliment. Our view is that judging men as more rational suggests, by implication, that they are detached, colder, and unsympathetic—all negative attributes.

The relevant results for this test are reported in Table OA4.1, where the single ratings of women and men as more emotional than rational are regressed on the language of interview. What we find is evidence that supports our proposed mechanism, i.e., "gender-blindness." More precisely, respondents interviewing in Estonian judged women to be significantly less emotional than those interviewing in Russian—an evaluation, we emphasize, that is uncontaminated by what they think of men, since these ratings were answered in randomized order. But when it comes to bilinguals' rating of men, there is no reliable relationship between interviewing in

Estonian and judgments of males on this stereotype dimension. We view these results as incompatible with the contention that perceiving greater gender parity comes at men's expense.

Table OA.4.1 The Effect of Genderless Language: Stereotype Ratings

	Model 1: Emotional women: single rating (OLS)	**Model 2:** Emotional men: single rating (OLS)
Estonian interview	-0.21** (0.09)	0.00 (0.08)
Constant	4.81*** (0.06)	3.46*** (0.06)
N	1,165	1,158

Note: Dependent variables are indicated in column headings. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$, two-tailed tests.

We reach a similar conclusion by exploring two other dependent variables that are more expressly political: *Female political recruitment* and *Female candidates: men's expense*. These questions were also asked independently, sequentially, and in random order. If the "gender-blindness" mechanism is correct, we should observe a positive and reliable coefficient for the first item—increased female political recruitment—but a negative and reliable coefficient for the second item—making room for female candidates by displacing men. Table OA.4.2 presents the raw results and Figure OA.4.1 graphs the relevant predicted probabilities of strong support for these proposals for both Estonian and Russian interviewees. While respondents interviewing in Estonian are more supportive of increasing the political recruitment of women to top political posts (predicted probability of strongly agreeing for respondents interviewing in Estonian: 0.28, in Russian: 0.23; FD = 0.05, CI: 0.01, 0.09), they are less supportive of increasing space for women on party candidate lists if this move comes at males' expense (predicted probability of strongly agreeing for respondents interviewing in Estonian: 0.14, in Russian: 0.18; FD = -0.04, CI: -0.07, -0.01). Thus, our results consistently show that the effect of speaking a gender-less tongue stems from a de-emphasis on distinctions between women and men, rather than promoting females or devaluing males.

Table OA.4.2 The Effect of Genderless Language: Parity at Men's Expense?

	Female political recruitment (Ordered Probit)	Female candidates: men's expense (Ordered probit)
Estonian interview	0.14**	-0.15**
	(0.06)	(0.06)
Constant	---	---
N	1,154	1,143

Note: Dependent variables are indicated in column headings. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$, two-tailed tests.

Figure OA.4.1 Probability of Supporting Increased Female Political Recruitment versus Political Parties Reserving Space for Women on Candidate Lists, Even if Men Are Excluded
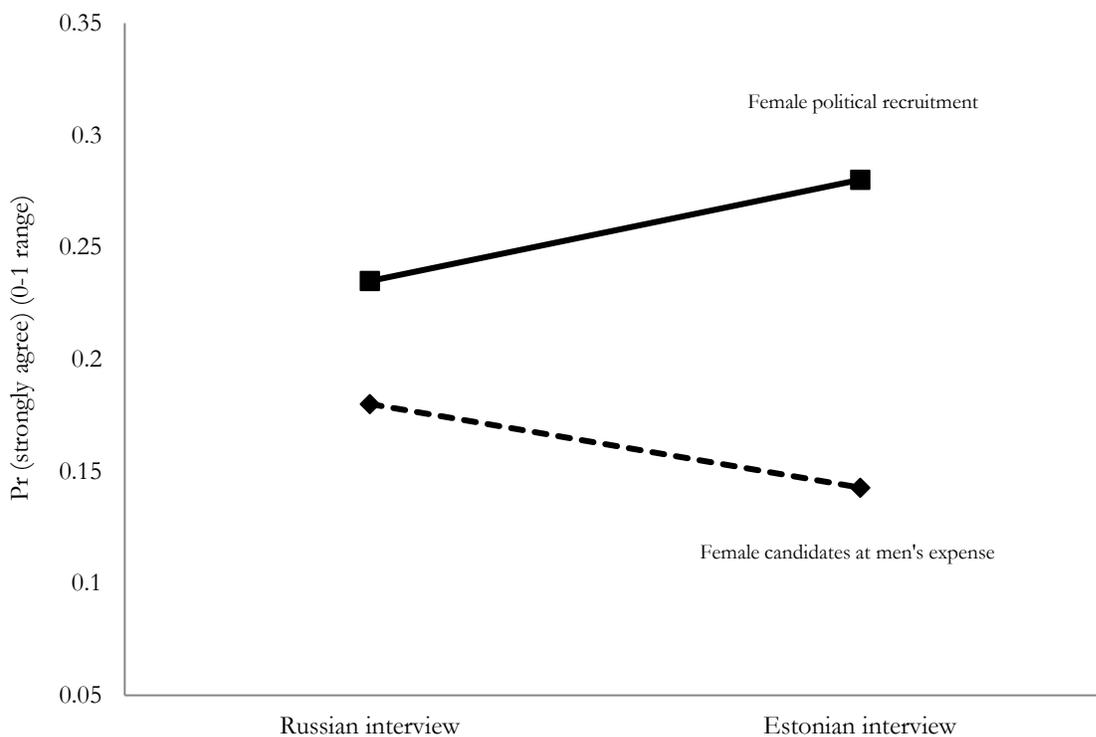
OA.5. Boundary Condition, Study 1

Table OA.5. Interview Language Has Null Effects on Items With Clearer Social Norms

	Buying sex	**Child suffers**	**Hit husband**	**Men better**
Estonian interview	-0.07 (0.08)	0.05 (0.07)	0.03 (0.08)	-0.06 (0.08)
Constant	0.34* (0.06)	0.09 (0.05)	-0.81* (0.06)	-0.31* (0.05)
N	1,071	1,130	1,160	1,1135

Note: Estimates are probit coefficients with standard errors in parentheses. *$p<0.05$, two-tailed tests

OA.6. Randomization and Balance Checks, Study 2

Table OA.6.1. Randomization Check: Language Manipulation is Unrelated to Pre-Treatment Covariates, Study 2

	Estonian Interview (=1)
Education	-0.033
	(0.165)
Female	-0.024
	(0.160)
Age	-0.013
	(0.006)
Russian first	0.149
	(0.276)
Prefer Russian	-0.209
	(0.292)
Constant	0.606
	(0.376)
N	1,200

Note: Labels in left-most column denote names of all pre-treatment covariates collected for Study 2. Entries are probit coefficients with standard errors in parentheses. An omnibus Wald test shows that one cannot reject the null that the covariates in Table OA.2.1 are simultaneously equal to zero, i.e., unassociated with the language manipulation (χ^2 = 5.07, Prob χ^2 (5) > 5.07 = 0.41, two-tailed).

The range and units of the covariates are as follows:

- *Education* is a dichotomous variable where '1' indicates R completed some college or higher and '0' indicates R completed secondary school or less.
- *Female* is a dummy variable where males are the omitted category.
- *Age* ranges from 22 to 76 in years.
- *Russian first* indicates bilingual R learned Russian first.
- *Prefer Russian* indicates bilingual R prefers to interview in Russian.

Table OA.6.2. Distributions of Pre-Treatment Covariates (Balance Check), Study 2

	Estonian interview	Russian interview	Chi-square test	Wilcoxon-Mann-Whitney test
Education (%)	41%	40%	$\chi^2 (1) = 0.017$ $p < 0.90$	---
Female (%)	53%	54%	$\chi^2 (1) = 0.060$ $p < 0.81$	
Age (median)	54	57	---	z = 1.968 Prob > |z| = 0.049
Russian first (%)	39%	37%	$\chi^2 (1) = 0.090$ $p < 0.764$	---
Prefer Russian (%)	31%	32%	$\chi^2 (1) = 0.007$ $p < 0.935$	---

Note: The differences in statistical tests performed reflect the nature of the covariates under analysis (e.g., dichotomous versus interval). All significance tests are two-tailed.

OA 7: Additional Information Related to Study 3

Variables included in the analysis

Women jobs (World Values Study item C001)
"Do you agree, disagree or neither agree nor disagree with the following statements?: "When jobs are scarce, men should have more right to a job than women""
1= disagree or neither agree nor disagree
0=agree

Women political leaders (D059)
"For each of the following statements I read out, can you tell me how strongly you agree or disagree with each. Do you strongly agree, agree, disagree, or strongly disagree?: "On the whole, men make better political leaders than women do""
1=strongly agreed
2=agreed
3=disagreed
4=strongly disagreed

University for girls (D060)
"For each of the following statements I read out, can you tell me how strongly you agree or disagree with each. Do you strongly agree, agree, disagree, or strongly disagree?: "University is more important for a boy than for a girl""
1=strongly agreed
2=agreed
3=disagreed
4=strongly disagreed

Women business executives (D078)
"For each of the following statements I read out, can you tell me how strongly you agree or disagree with each. Do you strongly agree, agree, disagree, or strongly disagree?: "On the whole, men make better business executives than women do""
1=strongly agreed
2=agreed
3=disagreed
4=strongly disagreed

Language spoken at home (G016)
This variable was used to identify respondent's language.

Sex (X001)
A binary variable indicating whether the respondent is female.

Age (X003)
Respondent's age in years.
Married (X007)

A binary variable indicating whether the respondent is married.

Education (X025)
Records the highest level of education for the respondent (8 categories).

Unemployed (X028)
A binary variable indicating whether the respondent is unemployed.

Income (X047)
Measures respondent's income decile.

Democracy is measured using the "polity2" variable from the Polity IV Project. The variable ranges from -10 (most authoritarian) to 10 (most democratic).

Genderless language is coded from the *World Atlas of Language Structures* (WALS) (Dryer and Haspelmath 2013), using their indicator "44A: Gender Distinctions in Independent Personal Pronouns." WALS codes each language into one of the following six categories: (1) in 3rd person + 1st and/or 2nd person; (2) 3rd person only, but also non-singular; (3) 3rd person singular only; (4) 1st or 2nd person but not 3rd; (5) 3rd person non-singular only; (6) No gender distinctions. The binary variable genderless language used in main text is coded "0" if 44A= "1" ("strongly gendered") and "1" ("non-gendered") otherwise. A categorical version of the same variable, referred to in footnote 27 of main text and used in the alternative analyses presented in Table SI.8.2 below was coded as follows: "2" if 44A = 6 ("non-gendered"); "1" if 44A = 2, 3, 4, or 5 ("weakly gendered"); "0" if 44A = 1 ("strongly gendered")

Table OA.7.1: Descriptive Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Women jobs	176,939	0.60	0.49	0	1
University for girls	172,827	2.96	0.92	1	4
Women business executives	100,373	2.62	0.97	1	4
Women political leaders	168,902	2.41	0.99	1	4
Gender equal attitudes	95,862	0.00	1.01	-2.16	1.62
Genderless language	180,447	0.63	0.48	0	1
Genderless language (categ.)					
Weakly gendered	180,447	0.43	0.49	0	1
Genderless	180,447	0.20	0.40	0	1
Sex	180,270	0.52	0.50	0	1
Age	180,067	41.28	16.36	15	99
Unemployed	174,585	0.09	0.28	0	1
Income	165,469	4.65	2.31	1	10
Education	168,054	4.75	2.23	1	8
Married	180,006	0.58	0.49	0	1
Democracy	174,642	4.29	6.41	-10	10

Note: The sample size varies significantly across the different models. These descriptive statistics were calculated from the cases used in the model with the largest N (*Women Jobs* Model 1 in Table 3 of the main text).

Table OA.7.2 The Effect of Genderless Language on Opinions Toward Gender Equality, World Values Survey 1995-2014, Categorical Coding of Genderedness

	Women Polit. Leader	Univ. for Girls	Women Busin. Exec.	Women Jobs	Gender Equal Attitudes
Weakly gendered	0.186***	0.227***	0.126*	0.388***	0.191**
language	(0.063)	(0.077)	(0.069)	(0.109)	(0.089)
Genderless	0.176**	0.409***	0.113**	0.534***	0.258***
language	(0.080)	(0.077)	(0.052)	(0.109)	(0.071)
Sex	0.281***	0.233***	0.323***	0.551***	0.357***
	(0.014)	(0.018)	(0.0217)	(0.041)	(0.024)
Age	-0.002***	-0.002***	-0.002***	-0.007***	-0.002***
	(0.000)	(0.001)	(0.001)	(0.002)	(0.001)
Unemployed	0.025**	-0.027***	0.014	-0.075**	-0.001
	(0.011)	(0.010)	(0.014)	(0.035)	(0.015)
Married	-0.026***	-0.009	-0.014	-0.153***	-0.030**
	(0.009)	(0.008)	(0.010)	(0.031)	(0.012)
Democracy	0.002	0.010	0.005	-0.010	-0.017
	(0.010)	(0.010)	(0.017)	(0.014)	(0.020)
Country FE	YES	YES	YES	YES	YES
Wave FE	YES	YES	YES	YES	YES
Education	YES	YES	YES	YES	YES
Income	YES	YES	YES	YES	YES
Constant	1.971***	2.432***	2.261***	-0.296	-0.359***
	(0.078)	(0.073)	(0.097)	(0.218)	(0.111)
N	135,827	138,761	83,324	141,632	79,937
N (countries)	84	84	71	83	71
R-squared	0.041	0.041	0.049		0.073

Note: Dependent variables are indicated in column headings. Robust country-clustered standard errors in parentheses. ***p<0.01, **p<0.05, *p<0.1