

Gender Dynamics in Elite Political Contexts: Evidence from Supreme Court Oral Arguments*

Bryce J. Dietrich[†] Ryan D. Enos[‡] Maya Sen[§]

First Draft: November 16, 2016

This Draft: August 27, 2017

Abstract

Significant attention has been devoted to understanding gender dynamics across a variety of contexts, including in boardrooms and in academia. We examine one place where unstructured conversation among elite male and female political actors takes place—the U.S. Supreme Court. We do so by investigating both verbal and non-verbal cues. Non-verbal cues, such as changing the tone of one's voice, not only structure interactions by signaling substantive content such as dominance, but also largely occur below conscious awareness. Looking at 33 years of Supreme Court oral arguments, we find, consistent with other studies on gender dynamics, that male Justices are more likely to talk over and to interrupt female Justices and female lawyers. Specifically, male Supreme Court Justices increase their vocal pitch when questioning, or being questioned by, female Justices and lawyers, a signal consistent with an unwillingness to yield the floor. We further present evidence that these emotionally unbalanced interactions have downstream consequences. Specifically, when female Justices and lawyers are talked over in this fashion, they then become more reticent to participate.

*Comments and suggestions welcome. We are grateful to Lee Epstein and Joseph Smith, and conference participants at MPSA for helpful feedback. Authors' names listed in alphabetical order.

[†]Department of Political Science, University of Iowa, 341 Schaeffer Hall, Iowa City, IA 52242 (bryce-dietrich@uiowa.edu, <http://www.brycejdietrich.com>).

[‡]Department of Government, Harvard University, 1637 Cambridge St., Cambridge, MA 02138 (renos@gov.harvard.edu, <http://ryandenos.com>).

[§]John F. Kennedy School of Government, Harvard University, 79 John F. Kennedy Street, Cambridge, MA 02138 (maya_sen@hks.harvard.edu, <http://scholar.harvard.edu/msen>).

1 Introduction

Well known among Supreme Court observers was the warm friendship between Justices Ruth Bader Ginsburg and Antonin Scalia outside the courtroom. Inside the courtroom, however, the two often interacted, but not always in positive ways. For example, in *Rodriguez v. United States*, one of the last interactions before Justice Scalia died in February of 2016, a calm Ginsburg was interrupted by an animated Justice Scalia roughly ten minutes in; Scalia then took over Ginsburg’s line of questioning and continued quizzing the attorney. Following this interaction, Justice Ginsburg lowered her tone in subsequent questioning and proceeded to ask fewer questions of the litigants (9) compared to her average (11.5).

As the illustration of Ginsburg and Scalia suggests, such interactions—even ones that do not appear on their face to be negative—can have downstream consequences on subsequent participation by female Justices. Understanding these interactions is, moreover, of substantive importance. Many papers have looked at how ordinary women and men communicate on topics of political importance generally (Karpowitz, Mendelberg and Shaker, 2012), finding that women participate less than men. But fewer papers have been able to examine how men and women interact in elite political environments, where important policy decisions are reached. Seeing gendered differences in elite communication may shape the contours of policies, regardless of women’s descriptive representation.

We engage the important substantive question of how men and women interact in elite political contexts by examining the U.S. Supreme Court. Unlike other elite political environments (for example, Congress), in which speech is often planned and scripted, the interactions between the Justices at oral argument are unstructured and fairly extemporaneous (Black et al., 2011). In addition, the Justices operate as co-equals, and, although seniority is extremely important, no political or professional competition structures the conservation, which some scholarship (e.g., Kanthak and Woon, 2015) has shown deters female participation. Lastly, also unlike other elite political environments, the Justices of the Supreme Court

interact with each other repeatedly over time, thereby providing a lens into how gendered interactions may have downstream consequences for women’s participation.

However, even institutions such as the Supreme Court have informal rules of decorum, which limits the information revealed by more manipulable verbal cues, such as interruptions. For that reason, we also examine *non-verbal cues*. Non-verbal cues, such as changing the tone of one’s voice, not only structure interactions by signaling substantive content such as dominance, but also largely occur below conscious awareness. In other words, Justices are reluctant to interrupt each other, but they pay less attention to subtle changes in the tone of their voice. This is why, in other contexts, non-verbal turn-taking has been shown to be an important part of gendered interactions in classrooms (Swann and Graddol, 1988) and student organizations (Rosenthal, Jones and Rosenthal, 2003). Our inquiry here concerns what similar non-verbal dynamics reveal about the interactions between men and women during oral arguments.

To address these issues, we tap into Supreme Court oral arguments audio, a dataset that encompasses 2,625 unique cases spanning over 2,544 hours of recordings. Processing this audio and analyzing the vocal inflections used by the Justices reveals important gender-based patterns. First, male Supreme Court Justices increase their vocal pitch when questioned by female Justices, a signal that is consistent with an unwillingness to yield the floor. Second, when immediately following female Justices, male Justices are much more likely to raise their vocal pitch, a signal they are attempting to seize the floor. Note that we find this behavior both in speech that male Justices direct toward other female Justices and also toward female lawyers making arguments before the Court. Both of these findings are consistent with existing scholarship suggesting that women are often “talked over” by their male counterparts.

Second, we examine the downstream consequences of these interactions on female Justices and female lawyers. Specifically, we examine the vocal intonations by women and their subsequent participation. Our first finding is that increased vocal pitch and interruptions by male Justices leads female Justices to be more reticent in asking questions during oral

arguments. Our second finding is that increased vocal pitch by male Justices leads female lawyers to become more disengaged. These results demonstrate even subtle gender dynamics can have a significant effect on the way women participate and, specifically, lead to emotional disengagement over time.

Our contributions are threefold. First, we contribute to a growing literature in political psychology documenting how women and men operate in mixed-gender environments, here leveraging a fairly unique elite context in the U.S. Supreme Court. Second, we find non-verbal cues are also an important part of gender dynamics on the Supreme Court and somewhat distinct to verbal cues. And lastly, we demonstrate that gender dynamics not only influence how Justices communicate but they can have important downstream effects on subsequent female participation.

This paper proceeds as follows. In Section 2, we discuss existing scholarship on gender dynamics among elites in other environments as well as on the Supreme Court. We use this literature to develop our key theoretical predictions regarding gender dynamics on the Court in Section 3. In Section 4, we explain our data, which involve over thirty years of U.S. Supreme Court oral arguments. We further explain how we used these data to extract both textual data as well as non-verbal cues, including the relative vocal pitch of the Justices. Our main results are presented in Section 5, which shows how Justices tend to “talk over” female Justices and lawyers. We also show how these patterns may encourage female Justices—and, to a lesser extent female lawyers—to withdraw from oral arguments, as evidenced by a reduction in vocal pitch after their turn is taken by a male Justice. We conclude in Section 6 with a summary and a discussion of avenues of future research. More details about our data, as well as additional results, can be found in the Supplemental Information.

2 Gender Dynamics, Vocal Pitch, and Supreme Court Oral Arguments

How are women's voices heard when it comes to mixed-gender environments? Recent scholarship has uncovered differences in how men and women communicate and participate in deliberation and discourse. The majority of this literature has found that women are less likely to speak in mixed-gender environments (Karpowitz, Mendelberg and Shaker, 2012), more likely to be interrupted or talked over by both men and women (Karpowitz and Mendelberg, 2014), and less likely than men to have confidence in their academic or professional abilities (e.g. Pomerantz, Altermatt and Saxon, 2002). These findings are robust and extend to a variety of contexts, including elite environments—such as corporate boardrooms (e.g., Baxter, 2014), legislative committees (Kathlene, 1994; Rosenthal, 2000), etc.—and also to more mundane settings—such as doctor's offices and day-to-day interactions between spouses.

The literature also makes clear that an important component of gender dynamics is the back and flow of conservation and, specifically, whether one speaker interrupts the other. When it comes to interruptions, several important studies, mostly in sociology and psychology have observed couples, pairs, or other mixed-gender groups in semi-structured discussion—some of these occur in a laboratory setting while others in a classroom setting. By and large, this literature has documented that men are more likely to interrupt women than they are other men (or vice versa) and that this is a demonstration of various forms of power and dominance. (For a review of this literature and meta-analysis, see Anderson and Leaper (1998).) For example, Hancock and Rubin (2015) paired male and female partners and asked them to have a conservation. They found no differences based on speaker gender, but did find that male speakers—when paired with a female speaker—were significantly more likely to interrupt. Another finding comes from Kollock, Blumstein and Schwartz (1985), who observe couples interacting about a conflicting “story,” finding that gender and power were closely correlated. Specifically, they find that the more powerful person in the couple

would interrupt more frequently and ask fewer follow-up questions.

More recently, this literature has been extended to the Supreme Court. Beginning with Johnson, Black and Wedeking (2009) and later Kimmel, Stewart and Schreckhise (2012), scholars have generally been interested in the degree to which Justices interrupt one another. Ultimately, Johnson, Black and Wedeking (2009) found that ideological differences were a significant predictor of interruptions, but they also found "interruptions are not as pervasive as media accounts, attorneys, and Justices indicate" (349-350). Most recently, Feldman and Gill (2017) extended this work to gendered interruptions. Using oral argument transcripts from 2004–2014, they find female Justices are not only more likely to be interrupted, but also tend to incorporate more polite language into their interactions.

We consider a different question, which is whether gender also structures the non-verbal interactions between Justices, specifically *changes in their vocal pitch*. We do so for several key reasons. The first is pragmatic: Justices are on the record as saying that they are self-conscious of when they speak and whom (and in what context) they interrupt, generally adhering to long-standing norms of collegiality and decorum. Indeed, while the U.S. Supreme Court does not have civility codified into their code of conduct similar to other courts,¹ many of the Justices have written about the need for civility (e.g., Thomas, 1996, 1997, 1998), so much so that when speaking to the American Law Institute, Chief Justice Warren E. Burger said, "to speak of the necessity for civility to a body noted for the civility of its debates may seem like speaking to The College of Cardinals about faith and morals" (Burger, 1975, 8). Justice Ruth Bader Ginsburg argued in another lecture that Justices actually try their best to wait for their colleagues to finish speaking:

A Justice at one end of our long, barely-curved, nine-member bench cannot always see what is going on at the other side. In particular, she may not notice

¹For example, the Supreme Courts of Delaware, Hawaii, and Pennsylvania all have some language prohibiting judges from using "hostile, demeaning or humiliating" language towards other judges during oral arguments. Similarly, several state Supreme Courts also ask judges to refrain from "unreasonable or unnecessary judicial interruption."

that an inquiry is emerging or continuing several seats away. When we interrupt or step on each other's questions, that is the reason why. We do not intentionally cut off a colleague. One tries to await another Justice's questions, but is not always successful. (I prove this from time to time.) (524)

Because of these norms, Justices may consciously avoid interruptions—at least to the best of their abilities. This suggests looking at interruptions exclusively may mask important patterns, particularly with regards to gender.

The second is substantive and speaks to the important role that non-verbal communication might have in elucidating inter-group patterns. For example, Ko, Sadler and Galinsky (2014) found vocal pitch was associated with a speaker's "rank." Here, they randomly assigned speakers to a "high rank" and "low rank" condition. In the former, speakers were told to imagine (a) they had a strong alternative offer, (b) valuable information, and (c) high status in their work place. Not only did high ranking speakers raise their vocal pitch, but respondents in two separate experiments correctly classified the high ranking speakers using their vocal pitch. Ko, Sadler and Galinsky (2014) also provides convincing anecdotal evidence that increases in vocal pitch are associated with increased status and perceived power. Before her election in 1979, Margaret Thatcher received voice training in order to "present a more powerful persona" (Ko, Sadler and Galinsky, 2014, 11). Using audio clips from Gallafent (2008), Ko, Sadler and Galinsky (2014) found Thatcher spoke at 178.88Hz before her voice training, while after she spoke at 198.75Hz – suggesting vocal pitch is associated with a speaker's perceived dominance.

Linguists have applied both sets of findings to conversation dynamics in a variety of settings, like political interviews (Beattie, Cutler and Pearson, 1982; Zellers, 2013; Zuraidah and Knowles, 2006) and group deliberations (Cutler and Pearson, 1986; Duncan and Fiske, 2015), ultimately finding lower vocal pitch is associated with turn transfers whereas higher vocal pitch is associated with turn taking. These findings follow from the well-documented relationship between vocal pitch and emotional activation (for review, see Mauss and Robin-

son, 2009). In order for turn-taking to occur, speakers must feel as though their contribution is more valuable than what is currently being said. Ultimately, this naturally puts them into an activated emotional state as they (a) become increasingly excited to make their next point and (b) ultimately must raise their tone in order to signal that excitement – and their intent to interrupt – to the current speaker. In both instances, their vocal pitch naturally increases as the rest of their muscles (including their vocal cords) naturally tighten as their body “gets ready for action” (Posner, Russell and Peterson, 2005).

For example, when male Justices address female Justices they likely feel as though they have something important to contribute to oral arguments. In these instances, emotional activation is more likely since Justices are probably more excited to speak when they have something important to say. When such activation occurs while Justices are speaking, it is indicative of an unwillingness to yield. The same relationship exists when Justices are trying to seize the floor. In the former, activation occurs because Justices feel as though what they are currently saying is more important than the next Justice’s contribution. In the latter, Justices feel as though what they need to say is more important than the words that are currently being uttered by another Justice. If male Justices are more likely to display these types of non-verbal signals towards female Justices and lawyers, then it suggests they generally feel as though their contributions to oral arguments are more important than what their female colleagues have to offer. Although we are unable to definitely say these types of non-verbal signals are indicative of implicit gender bias, changes in vocal pitch occur largely subconsciously (Ekman et al., 1991), implying non-verbal signals may be an honest indicator of the underlying gender dynamics on the Supreme Court.

Why do gender dynamics matter? First, several papers point to tentative conclusions that suggest more aggressive interactions—as shown by higher vocal pitch—could curtail or discourage subsequent participation of female Justices. For example, Johnson, Black and Wedeking (2009) and Feldman and Gill (2017) argue that interruptions may “cause Justices to become reticent to ask questions,” suggesting aggressive turn-taking, such as increases in

vocal pitch, may have important downstream consequences. Supporting evidence is found in jury deliberations. Beginning with Strodtbeck and Mann (1956), scholars have found men are not only more likely to dominate jury deliberations (for review, see Fowler, 2005; Marder, 1987; Toro, 2015), but they are also more likely to serve as the foreperson (Beckham and Aronson, 1978). Although it is difficult to determine whether such dominant behavior is associated with male jurors raising the tone of their voice, Nemeth, Endicott and Wachtler (1976) found male jurors were rated as being more active and aggressive, traits typically associated with increases in vocal pitch (see Mauss and Robinson, 2009). If female jurors withdraw from deliberations in response to non-verbal turn-taking, then it suggests female Justices and lawyers may “become more inhibited” (Feldman and Gill, 2017, 62) after an aggressive interaction with a male Justice.

Second and relatedly, Boyd, Epstein and Martin (2010) argue that female judges “possess unique and valuable information emanating from shared professional experiences” (391). Looking at the U.S. Supreme Court, if male Justices are “talking over” female Justices because they disagree with what they are saying, then male Justices may not perceive them as being persuasive, even on issues where they are expected to have some expertise (Boyd, Epstein and Martin, 2010, 392). Perhaps more importantly, if male Justices are “talking over” female Justices because they agree with them, then we are highlighting a possible negative consequence of Boyd, Epstein and Martin (2010)’s “informational” approach. Indeed, while we agree “the greater the diversity of participation by [judges] of different backgrounds and experiences, the greater the range of ideas and information contributed to the institutional process” is well-founded, our study demonstrates that it may be incredibly difficult for women to make informational contributions even when men agree with what they are saying (Boyd, Epstein and Martin 2010, 406-407; Epstein, Knight and Martin 2003, 944).

3 Theoretical Expectations Regarding Gender Dynamics on the Supreme Court

Motivated by this literature, we now turn to develop some concrete observable implications. As we have noted, scholars to date have focused the majority of their attention on interruptions, especially when considering the Supreme Court. Although subtler, non-verbal turn-taking is much more common during oral arguments and also guided in part by subconscious reactions. Nonetheless, we can take previous scholarship on verbal turn-taking (interruptions and other analyses of interactions) inside and outside the Supreme Court in formulating several key hypotheses.

Hypothesis 1a: Male Justices will raise their vocal pitch more towards female Justices than male Justices. Non-verbal cues are often used to facilitate turn-taking. For example, when speakers do not want to yield, their vocal pitch increases as they attempt to signal to those around them they are still committed to the point they are trying to advance. Similarly, when interlocutors want to usurp another speaker's turn they increase their vocal pitch to demonstrate they are particularly invested in either the question or statement they want to interject.

Based on previous literature, we expect male Justices likely behave differently towards female Justices, but how that affects their non-verbal behavior is less clear. If male and female Justices exhibit traditional gender roles, then we expect that male Justices will tend to raise their vocal pitch when addressing a female Justice. When they are the ones asking questions or making statements, such increases signal to the female Justice their unwillingness to yield. When responding to a question or statement made by a female Justice an increase in vocal pitch signals their willingness to end the female Justice's turn.

Hypothesis 1b: Male Justices will raise their vocal pitch more towards female lawyers than toward male lawyers. Gender dynamics also likely influence the interac-

tions between male Justices and female lawyers in a similar manner. If male Justices and female lawyers are exhibiting well-established gender roles, then male Justices should raise their vocal pitch when addressing female lawyers. When asking questions or making statements directed towards female lawyers, such increases signal a male Justice's unwillingness to yield. When responding to arguments advanced by female lawyers, these same increases indicate a willingness of a male Justice to end a female lawyer's turn.

Hypothesis 2a: Female Justices will withdraw from oral arguments after male Justices raise their vocal pitch After being "talked over," we expect female Justices are less excited to contribute new information to oral arguments. Indeed, when female Justices are "talked over" by their ideological allies, then they likely see little point in contributing new information to oral arguments since their male colleagues not only have similar opinions but seem to hold them more fervently – as indicated by their increased vocal pitch. Similarly, when "talked over" by their ideological opponents, then female Justices may decide to reserve their opinions for conference or personal communications where the emotional tenor is often less combative. Regardless, female Justices are likely to become more disengaged after being "talked over" by their male colleagues – an effect we ultimately isolate using their own vocal pitch.

Hypothesis 2b: Female lawyers will withdraw from oral arguments after male Justices raise their vocal pitch. Although we expect to find a similar pattern for female lawyers, we also expect these relationships to be less pronounced. Unlike female Justices whose participation is voluntary, female lawyers have to ask and answer questions, regardless of the Justices' demeanor. This gives female lawyers a strong incentive to carry on, even when male Justices are consistently "talking over" them. Thus, we expect female lawyers will withdraw from oral arguments when male Justices raise their vocal pitch, but changes in the vocal pitch of male Justices will have a greater effect on female Justices as compared

to female lawyers.

Hypothesis 3: Interruptions. For all of the above hypotheses, we provide interruptions as a point of comparison. Both verbal and non-verbal signals can be used to either retain or take control of the floor. Interruptions are a very good example of the former, we offer vocal pitch as an example of the latter. Even though there is considerably more variance in vocal pitch, we expect both measures should be largely consistent. We ultimately find they are, suggesting both are important parts of turn-taking behavior on the Supreme Court.

4 Description of Supreme Court Oral Arguments and Data

To give some brief background, the Court hears arguments from the first Monday in October until sometime in June, hearing approximately 70–80 cases per year. The arguments last about an hour (usually with 30 minutes per side) and roughly follow the same format. First, each lawyer begins by formally addressing the Chief Justice and the Court and providing opening remarks. At this point, he or she will usually (but not always) be interrupted with questions and commentary from the Justices. Such questions can be brief interjections or longer, multi-pronged questions. In addition, it is not uncommon for the Justices to address points made by other Justices, and to have lengthy back and forths with relatively little input from the lawyers themselves. Importantly, there are no “rules” governing the order of speech, nor how the discussion ought to be structured. However, it is usually the case that the Justices are the ones to interrupt attorneys, and not vice versa.²

Our data on these arguments come from three related sources, which we then combine to create our final dataset. The first are transcript data from the Supreme Court oral arguments, which we obtained from the Oyez Project at the Chicago-Kent College of Law. We collected

²The official Supreme Court “Guide for Counsel” recommends that attorneys “[n]ever interrupt a Justice who is addressing you...If you are speaking and a Justice interrupts you, cease talking immediately and listen.”

data for 33 terms, from 1982 to 2014; leaving us with data for 4,840 unique cases. We used the transcript data to extract several variables of substantive importance. First, we added each speaker’s “type,” specifically whether the speaker was a “Justice” or “advocate.”³ Second, we added the description for each advocate to identify which advocate represented the petitioner or the respondent. We further used the transcript data to assist in parsing discrete utterances and to pinpoint speaker dyads. For example, the transcript provided information on whether Justice Antonin Scalia was addressing the petitioner advocate, the respondent advocate, or another Justice (and who that Justice was). We also used the transcript data to identify instances where speakers interrupted each other; these are denoted in the transcript with utterances that ended in the use of en and em dashes (“-” and “—”). By contrast, utterances that were uninterrupted ended in standard punctuation (for example, a period or a question mark).

Second, part of our hypotheses relies on differing interactions between men and women, and how these interactions could be manifested in emotions and subconscious reactions—and not just interruptions, which are easier for people to temper consciously. To explore this, our first data source was therefore oral arguments audio recordings, which have been recorded since the installation of recording equipment in October of 1955. We used audio recordings maintained by the Oyez Project. We next divided the audio recordings by speaker—both Justices and also by each of the speaking attorneys—to get discrete utterances. This left us with 442,106 discrete utterances, 146,335 of which were spoken by Justices, 149,789 by petitioner attorneys, and 145,981 by respondent attorneys.

We then took each utterance and processed it to obtain a measure of emotional arousal. For this, we looked to vocal pitch, as noted above. To estimate the level of vocal pitch in each utterance, we processed the recordings using *Praat*.⁴ This left us with a measure of

³We discarded utterances made by two other types of speakers—“unidentified” and “other.”

⁴This commonly used software (<http://www.fon.hum.uva.nl/praat>) implements the algorithm outlined by Boersma (1993). Similar to other algorithms that focus on time-domain periodicity, *Praat* estimates the fundamental frequency by dividing the autocorrelation of a windowed signal by the autocorrelation of the window itself. One must assume the signal is stationary within each window, which is why the algorithm

vocal pitch for each utterance. Aggregating up to the Justice level also provides a useful measure of Justice-level vocal pitch, which we can then use to normalize each individual utterance. The Justices' baseline vocal pitch is reported in Table 1. Of the female Justices, Sandra Day O'Connor had the highest average vocal pitch (196.28Hz) and had the highest standard deviation (31.87Hz). For the male Justices, Thurgood Marshall had the highest average vocal pitch (147.84Hz), whereas Harry Blackmun had the highest standard deviation (46.96Hz).

Pairing the audio data with the transcript data revealed other patterns of note. For example, on average, each question and non-question was 15 and 21 seconds long, respectively. William Burger asked the longest question which was 20 minutes long. The shortest question, asked by Antonin Scalia, was less than a second. For non-questions, Scalia gave both the longest and shortest. The former was close to four minutes long. The latter was less than a second. For reasons explained below, these incredibly short questions and non-questions were excluded from the analysis, although they do not substantively affect the results. Ultimately, we found William Brennan spoke the least, followed by Clarence Thomas. In total, Brennan spoke for a little under 3 minutes across 14 questions/non-questions. Clarence Thomas spoke for a little over 11 minutes across 65 question/non-questions. On average, Justices spoke a little over 25 hours across 7,702 question/non-question, suggesting Brennan and Thomas barely spoke. Antonin Scalia was the most loquacious Justice, asking 12,382 questions and making 15,523 statements. In total, Scalia spoke for 79.82 hours, which is over three times as much as the average Justice. His next closest competitor, Stephen Breyer, spoke close to 7 hours longer than Scalia, but did so using 3,288 fewer questions and 4,880 fewer non-questions. The differences found between Scalia and Brennan should be taken into

divides the audio file into small segments (around 60ms), then takes the average. Specifically, to use this software, one has to set five parameters: the pitch floor, pitch ceiling, window length, window shape, and voicing threshold. We set the pitch floor and ceiling to 50Hz and 300Hz, respectively. This resulted in a window length of 60ms. For both the window shape and voicing threshold we used the default settings. Additional details are in the Supplemental Information. To ensure reliability, we excluded any questions and non-questions that were less than one second.

Table 1: Average vocal pitch and standard deviation for U.S. Supreme Court Justices, 1982–2014.

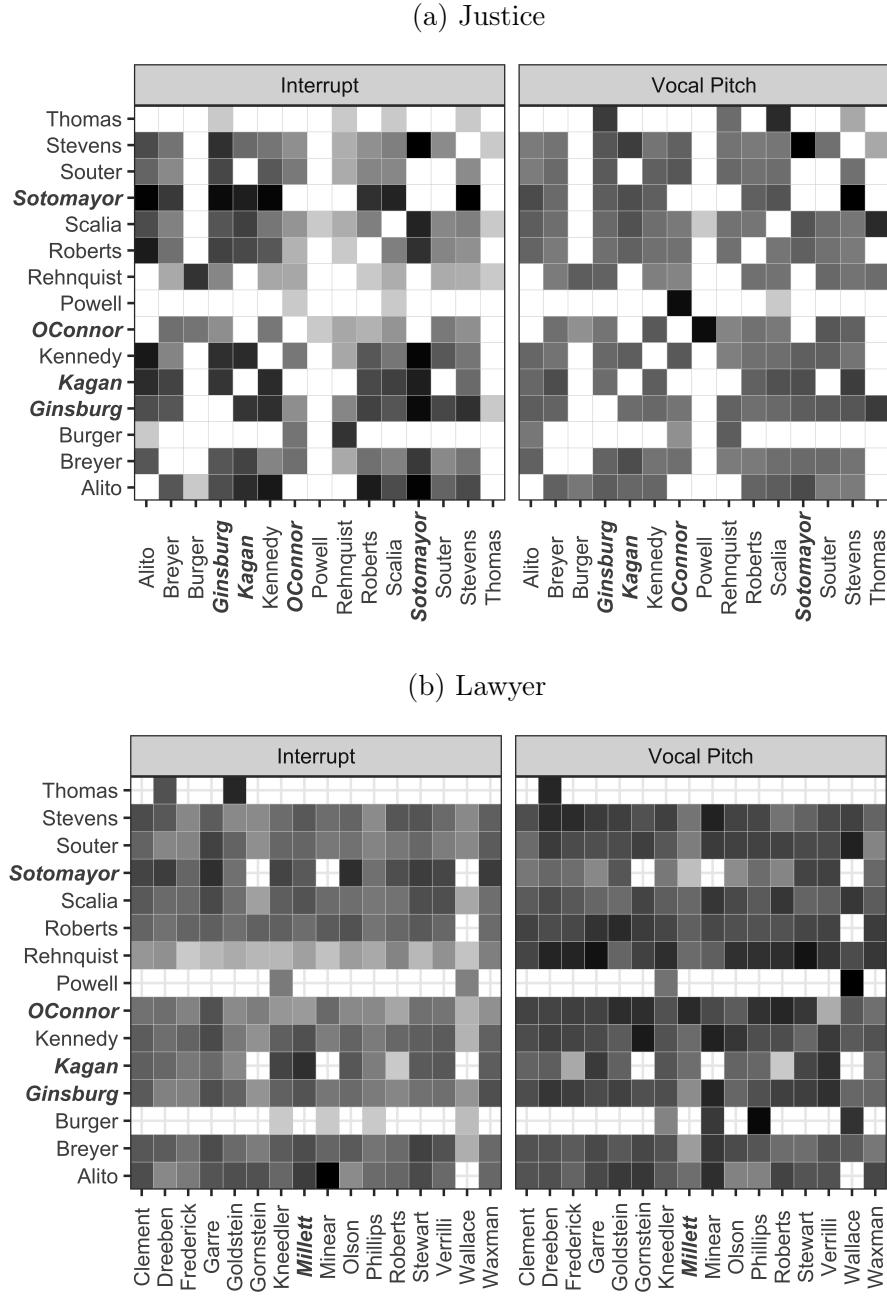
Justice	Pitch Mean	Pitch SD	Questions	Non-Questions	Total
Sandra Day O'Connor	196.28	31.87	3,505	3,115	6,620
Elena Kagan	171.60	28.05	1,616	1,363	2,979
Sonia Sotomayor	166.14	25.07	3,345	3,165	6,510
Ruth Bader Ginsburg	163.82	22.54	7,284	8,509	15,793
Thurgood Marshall	147.84	33.75	292	173	465
Antonin Scalia	136.57	31.97	12,382	15,523	27,905
David Souter	135.62	29.43	5,404	5,816	11,220
John Roberts	133.07	28.24	5,261	9,308	14,569
William Brennan	127.48	28.63	5	9	14
Stephen Breyer	126.95	30.62	9,094	10,643	19,737
Harry Blackmun	122.24	46.96	63	31	94
Anthony Kennedy	121.52	22.31	6,208	7,191	13,399
William Rehnquist	121.07	25.25	3,098	5,659	8,757
Warren Burger	119.65	33.49	539	1,279	1,818
Lewis Powell	117.01	36.97	78	42	120
Byron White	116.58	33.42	193	240	433
John Paul Stevens	116.21	24.74	5,505	5,727	11,232
Samuel Alito	112.66	24.67	3,366	1,239	4,605
Clarence Thomas	101.81	24.30	31	34	65
Average	134.43	29.59	3,540	4,161	7,702

Note: Measurements of vocal pitch are in Hertz (Hz). To calculate each we used all of the audio from Justices' questions and non-questions. The number of questions and non-questions are reported in the corresponding columns. In the last column, we report the total number of utterances. Averages of each column are reported at the bottom of the table. The table is sorted by the mean vocal pitch.

consideration when evaluating our results. The same can be said for Breyer and Thomas.

Figure 1 plots the dyadic relationships between the Justices (Panel A) and lawyers (Panel B). In all plots we scaled each square to range from the minimum (white) to maximum (black) for either the average vocal pitch or the interruption rate. In Panel A, we consider only Justice-Justice interactions, meaning Justices on the x-axis are speaking to Justices on the y-axis. Lawyer-Justice interactions are considered in Panel B. In total, we have 2,825 lawyers. To represent the data graphically, we only considered lawyers who argued at least 20 cases in front of the Court. We restricted the Justices appearing in the y-axis to only those that interacted with these lawyers. To make the plots comparable, in Panel A we

Figure 1: Justice-Justice and Lawyer-Justice Vocal Pitch and Interruption Rates



Note: Squares are scaled to the minimum (□) and maximum (■) of each variable. **Bolded** names are female. In Panel A, the x and y axes corresponds to the first Justice and second Justice. In Panel B, the x and y axes are the same, but we only consider lawyers addressing Justices. Given the number of lawyers in the data set, we restricted Panel B to lawyers who appeared in at least 20 cases. The corresponding Justices are the ones that had at least one interaction with those lawyers. To make the y axes consistent, we restricted Panel A to those Justices.

used these same Justices. Since we are interested in male-female interactions, we also bolded the names of female Justices and attorneys. Information about the remaining lawyers and Justices as well as other descriptive statistics about dyads not shown here can be found in the Supplemental Information.

A cursory look reveals that there are considerably more lawyer-Justice interactions. Indeed, there are 143 unique Justice-Justice dyads, whereas there are 12,309 unique lawyer-Justice dyads. Of the former, the maximum dyad is Stephen Breyer and Antonin Scalia who interacted 543 times. The smallest dyad was a single interaction which occurred in 23 dyads involving a number of Justices. For example, we only have three dyads that contain Clarence Thomas. Undoubtedly, we have more data from lawyer-Justice interactions. Here, Carter G. Phillips appears the most. He interacted with Stephen Breyer, Antonin Scalia, and Ruth Bader Ginsburg 513, 415, and 387 times in our data, respectively. This is not too surprising since he clerked for Chief Justice Burger, served as Solicitor General, and argued 61 cases in our data. As with the Justices, our smallest lawyer-Justice dyad only has a single interaction. We have 1,869 dyads of these dyads – again from a variety of Justices.

In terms of vocal pitch, the most contentious dyad involving a female Justice was Justice Stevens and Sotomayor who interacted the most in 2009 during *Bloate v. United States*, *Maryland v. Shatzer*, and *NRG Power Marketing v. Maine Public Utilities Commission*. In these interactions, Justice Stevens spoke over 2 standard deviations above his baseline, whereas Justice Sotomayor spoke essentially at her baseline – suggesting this dyad was a mostly one-sided affair. In Panel B, the only female lawyer to argue at least 20 cases was Patricia A. Millett. Her most contentious dyad involved Justice O'Connor whom she interacted with 45 times between 1998-2004. In these interactions, Justice O'Connor spoke over 0.555 standard deviations higher than Ms. Millett, suggesting not only was this a one-sided interaction, but attorneys may generally be reluctant to raise their vocal pitch when addressing Justices. Indeed, when lawyers are addressing Justices they speak essentially at their baseline, whereas when Justices speak to other Justices they tend to speak 0.43

standard deviations above their baseline.

Unfortunately, Figure 1 gives us little insights into the broader gender dynamics. Given that, we made some additional calculations. For example, when female Justices are addressing male Justices the average vocal pitch of the response is 0.29 – meaning male Justices respond on average 0.29 standard deviations above their baseline. This is significantly ($p < 0.001$) higher than the average vocal pitch of the response when female Justices are addressing each other. In these instances, female Justices respond with an average vocal pitch of 0.07 standard deviations which is over four times *less*. A similar relationship is found for female lawyers who receive a vocal pitch response of 0.15 standard deviations higher when addressing male Justices as opposed to female Justices. This difference is also statistically significant at 0.001-level. Although far from definitive, these initial comparisons suggest male Justices respond differently to women regardless of whether they are colleagues or attorneys arguing to the Court.

Lastly, several of our theoretical expectations engage the substance of the court cases involved and whether patterns observed could vary across and between different ideological coalitions. We therefore look to case-level and Justice-level covariates, which we take from the third data source, the Supreme Court Database, at Washington University-St. Louis (Spaeth et al., 2015). Specifically, we used case-level data on (1) the substantive issue area of the case, especially whether it involved sex discrimination, and (2) whether the lower court decision was in a liberal or conservative direction. (We linked this data to the oral arguments audio data by using the docket number of each case.) We also pair up each utterance with additional Justice-specific covariates. These are the (3) Justice's term-specific ideological score, as estimated by Martin-Quinn. (The Martin-Quinn scores are a continuous measure from conservative to liberal and have been substantially cross-validated with other measures.) Because speaking patterns might be influenced—implicitly or, perhaps, explicitly—by Justices' relative seniority, we also include a variable that captures (4) their seniority in each term. For lawyers, we calculated the number of years they appeared

in front of the Court. Along these same lines, we created binary variables indicating whether the questioning or answering Justice was Chief Justice at the time (5). To control for the ebb and flow of emotionality that can take place during oral arguments, we also (6) controlled for when the question began (in minutes).

5 Results

5.1 Male Justices Raise Their Vocal Pitch When Addressing Female Justices

We now turn to exploring the gender dynamics within the context of oral arguments. However, modeling interactions during oral arguments is challenging, since by definition the units of analysis are not independent from one another. Scholars modeling country interactions overcome a similar problem using country dyads (i.e., US-Russia). We apply the same strategy to modeling interactions during oral arguments, but instead of talking about one country going to war or imposing sanctions on another, we are interested in whether some Justices raise their vocal pitch when interacting with other Justices or lawyers.

We begin by examining the interaction of the Justices and the lawyers arguing the cases at oral argument. Recall that at each oral argument the Justices may interrupt the speaking attorney to ask questions, request clarifications, or to make observations. Because the lawyers tend to be extremely differential to the Justices, and because of the substantial power imbalance between the speaking attorneys and the Justices, these interactions lack the quality of the interlocutors being coequals, as is often the case in interactions in elite institutions. However, these interactions are also important and have the attractive quality of a large number of dyads: 314 female attorneys that participated in oral arguments between 1982 and 2014 and 2,511 male attorneys that participated during that same time period.

To better understand the models we estimated, consider Table 2. In these models, each observation represents an utterance dyad, or an instance where an utterance made by one speaker was followed by an utterance by another. This temporal progression is denoted by

the t . For example, $Male_t, Female_{t+1}$ refers to an utterance dyad where a male speaker (denoted by $Male_t$) is followed by a female speaker ($Female_{t+1}$). The dependent variable is the vocal pitch of the questioning male Justice scaled to standard deviations above and below his baseline. In these models, positive values imply the male Justice raised his vocal pitch when addressing the female attorney. In order to address the likely correlation across observations, we estimated multilevel linear regressions with a randomly varying intercept for each dyad.

Results for Justices' interactions with lawyers are in Columns 1 and 2. These models show that male Justices tend to lower their vocal pitch when addressing male attorneys, but when they are addressing female attorneys they tend to raise their vocal pitch. More specifically, when male Justices are addressing female attorneys as opposed to male attorneys their vocal pitch is 0.14 standard deviations higher. This result holds when additional controls are added to the model, including whether the questioning speaker is the Chief Justice, the seniority difference between the questioning Justice and speaker, and the ideology of the questioning Justice. In order to gain some leverage on the ideology of the argument being advanced, we also included a dummy variable capturing whether the lower court issued a liberal ruling. Even with these controls, there is a statistically significant ($p < .05$) difference between the interactions of male Justices and female attorneys, as compared to all other combinations.

While these results show a very real disparity in the tone of interactions when speaking with men or women before the Court, one might be more concerned with the interactions between Justices—after all, these are interactions taking place under strong norms of equality and between people at the peak of their profession. As such, it is a good test for the power of the gender dynamics observed in juries and other institutions to influence all levels of interactions, even in the most rarefied of settings.

Looking at Columns 3 and 4 of Table 2, we find that male Justices behave differently towards female Justices. Specifically, when a male Justice is addressing a female Justice the male Justice speaks with an average vocal pitch 1.11 standard deviations above his

Table 2: Male Justices Raise Their Vocal Pitch When Questioning Female Justices and Lawyers

(a) Lawyers

(b) Justices

	<i>Dependent variable:</i> Vocal Pitch _t		<i>Dependent variable:</i> Vocal Pitch _t	
	(1)	(2)	(3)	(4)
Fixed Effects				
Constant	-0.074*** (0.005)	-0.098*** (0.009)	Constant	0.372*** (0.049)
Male _t , Female _{t+1}	0.135*** (0.015)	0.131*** (0.015)	Male _t , Female _{t+1}	0.739*** (0.105)
Minutes _t		0.000 (0.000)	Minutes _t	-0.003** (0.001)
Ideology _{t+1}		-0.014*** (0.003)	Ideology _{t+1} - Ideology _t	-0.073*** (0.028)
Liberal Lower Court Decision		-0.016*** (0.006)	Seniority _{t+1} -Seniority _t	0.002 (0.004)
Ideology _{t+1} × Liberal Lower Court Decision		0.004 (0.003)	Chief Justice _t	-0.180 (0.132)
Seniority _{t+1} -Seniority _t		0.002*** (0.000)	Chief Justice _{t+1}	-0.101 (0.132)
Chief Justice _t		0.034*** (0.013)	Sex Discrimination	-0.285* (0.159)
Sex Discrimination		-0.023 (0.027)	Random Effects	
Random Effects				
Dyad	0.108* (0.328)	0.107* (0.328)	Dyad	0.164* (0.405)
N ₁	136,486	132,440	N ₁	7,351
N ₂	12,108	11,824	N ₂	142
Log Likelihood	-183,691.300	-178,185.400	Log Likelihood	-13,218.230
AIC	367,390.600	356,392.900	AIC	-11,663.310
				26,444.470
				23,346.630

Note: We use the subscripts t and $t + 1$ to indicate the speaker order, meaning Justice_t is speaking to Justice_{t+1}. Male_t, Female_{t+1} indicates a male Justice is speaking to a female Justice. In all models the dependent variable is the vocal pitch of the questioning Justice (t) in standard deviations above his or her baseline. We view each utterance as an indicator of a latent trait that exists between Justices. Given that, estimates are from multilevel linear regressions with random intercepts for each dyad. Our dyadic modeling strategy follows directly from models of international relations in which the unit of analysis is often country dyads (e.g., US-Russia). In Models 1 and 2, Seniority_{t+1} is the number of years Attorney_{t+1} has appeared in front of the Court. Seniority_t is the number of years Justice_t has served. In Models 3 and 4, Seniority_t and Seniority_{t+1} are the number of years Justice_t and Justice_{t+1} have served, respectively. Levels of significance are reported as follows: * $p < .1$; ** $p < .05$; *** $p < .01$. Standard errors are reported in parentheses.

Figure 2: Predicted Values for Table 2



Note: This plot uses the coefficients from the first and third models in Tables 3 and 4. \square and \blacksquare imply Female_t , Male_{t+1} was set to 0 and 1, respectively. Error bars represent 95 confidence intervals from 1,000 bootstrapped samples. The Supplemental Information contains more information regarding the bootstrapping procedure.

baseline. Conversely, when a male Justice is addressing another male Justice his vocal pitch is approximately 3 times *less* (0.37). Similar results are found in Model 2. Here, when all other variables are held constant at their mean and modal values and Male_t , Female_{t+1} is set to 1, the predicted vocal pitch of the male Justice is 1.20 standard deviations above his baseline. This is again close to three times higher than the predicted vocal pitch when Male_t , Male_{t+1} is set to 0 (0.42).

Regardless of whether a male Justice is addressing a female Justice or attorney, their non-verbal behavior indicates they are unwilling to yield. More specifically, decreases in vocal pitch indicate a speaker is no longer committed to the position they are advancing, whereas increases in vocal pitch signal he or she does not want to give up their speaking turn. Male Justices consistently raise their vocal pitch when speaking before female Justices and attorneys, ultimately indicating they may be reluctant to yield the floor to women participating in oral arguments. Although this gender dynamic is undoubtedly important, it is difficult to compare to turn-taking signals, such as interruptions, where someone is

attempting to seize as opposed to retain a speaking turn. In these instances, interlocutors may also use increases in vocal pitch to demonstrate to the current speaker they intend to take the floor. If male Justices are generally unwilling to yield to female Justices, then we should also expect them to raise their vocal pitch when responding to female Justices. Table 3 tests this hypothesis.

In these models, we include the same variables with one notable exception. Instead of including a binary variable isolating dyads where male Justices are addressing female Justices ($Male_t, Female_{t+1}$), we included a variable isolating dyads where male Justices were responding to female Justices ($Female_t, Male_{t+1}$). As before, the temporal progress is denoted by the t , meaning when this variable equals 1 a male Justice speaks immediately after a female Justice. The dependent variable in the first two columns of Table 3 is the vocal pitch of the responding male Justice scaled to standard deviations above and below his baseline. In these models, positive values imply the male Justice raised his vocal pitch when answering the female Justice. The dependent variable in the third and fourth columns is whether an interruption occurred. In these models, a one indicates the male Justice interrupted the female Justice while she was addressing him with either a question or statement. Since we are interested in understanding whether the vocal pitch and interruption models yield similar results, we plotted the predicted values of Models 1 and 3. These can be found in the “Justice” panels of Figure 3.

Similar to Table 2, Models 1 and 2 report the estimates from a multilevel linear regression with a randomly varying intercept for each dyad. In Models 3 and 4, we include the same random intercepts, but given the dichotomous nature of our variable we estimated a multilevel logistic regression instead. Beginning with Model 1, we find that male Justices behave differently towards female Justices. Specifically, when a female Justice is addressing a male Justice the male Justice on average responds with a vocal pitch that is 0.34 standard deviations above his baseline. Conversely, when a female Justice is addressing another female Justice the responding Justice’s vocal pitch is 0.05 standard deviations – approximately

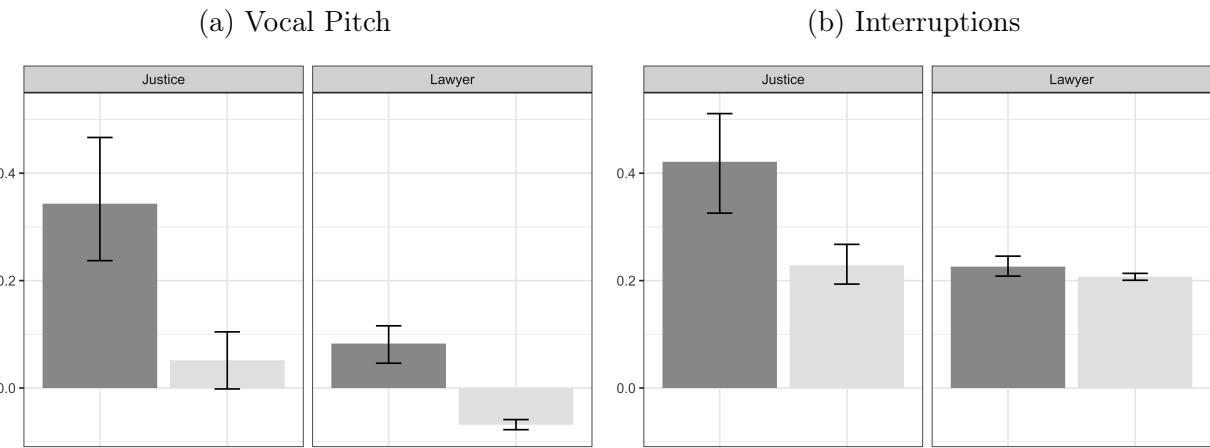
Table 3: Male Justices Raise Their Vocal Pitch When Answering Female Justices

	<i>Dependent variable:</i>			
	Vocal Pitch _{t+1}		Interrupted _{t+1}	
	(1)	(2)	(3)	(4)
Fixed Effects				
Constant	0.051* (0.027)	0.136** (0.054)	−1.173*** (0.112)	0.198 (0.197)
Female _t , Male _{t+1}	0.290*** (0.057)	0.247*** (0.060)	0.818*** (0.233)	0.493** (0.251)
Minutes _t		0.0005 (0.001)		−0.003** (0.002)
Ideology _{t+1} - Ideology _t		−0.029* (0.015)		−0.369*** (0.058)
Seniority _{t+1} -Seniority _t		0.001 (0.002)		0.028*** (0.008)
Chief Justice _t		−0.123* (0.070)		−0.883*** (0.293)
Chief Justice _{t+1}		−0.163** (0.067)		−0.522* (0.282)
Sex Discrimination		−0.118 (0.119)		0.037 (0.239)
Random Effects				
Dyad	0.038*** (0.195)	0.029*** (0.195)	0.989*** (0.995)	0.906*** (0.952)
N ₁	7,790	6,784	8,279	7,225
N ₂	143	132	143	132
Log Likelihood	−11,913.910	−10,562.600	−4,462.896	−3,879.068
AIC	23,835.830	21,145.190	8,931.791	7,776.137

Note: We use the subscripts t and $t + 1$ to indicate the speaker order, meaning Justice_t is speaking to Justice_{t+1}. Female_t, Male_{t+1} indicates a female Justice is speaking to a male Justice. In Models 1 and 2, we estimate a multilevel logistic regression given the outcome is whether the female Justice was interrupted (Interrupted_t). In Models 3 and 4, we estimated a multilevel linear regression since the outcome is the vocal pitch (in standard deviations) of the responding Justice ($t + 1$). Unit of analysis is each utterance with a random intercept for each dyad. Our dyadic modeling strategy follows directly from models of international relations in which the unit of analysis is often country dyads (e.g., US-Russia). Seniority_t and Seniority_{t+1} are the number of years Justice_t and Justice_{t+1} have served, respectively. Levels of significance are reported as follows: * $p < .1$; ** $p < .05$; *** $p < .01$. Standard errors are reported in parentheses.

seven times *less*. Not only is this a statistically significant relationship, but Figure 3 shows the difference is substantial. Although not plotted, similar results are found in Model 2. Here, when all other variables are held at their mean and modal values and Female_t , Male_{t+1} is set to 1, the predicted vocal pitch of the responding Justices is 0.32 standard deviations above their mean. This is nearly five times higher than the predicted responding pitch (0.07) when Female_t , Male_{t+1} is set to 0.

Figure 3: Predicted Values for Tables 3 and 4



Note: This plot uses the coefficients from the first and third models in Tables 3 and 4. \square and \blacksquare imply Female_t , Male_{t+1} was set to 0 and 1, respectively. Error bars represent 95 confidence intervals from 1,000 bootstrapped samples. The Supplemental Information contains more information regarding the bootstrapping procedure.

Models 3 and 4 yield similar results for interruptions. However, the effect size is somewhat smaller. Beginning with Model 3, when female Justices are addressing male Justices there is a 41 percent chance they will be interrupted. This is 11 percentage points higher than the predicted interruption rate when they are addressing a female Justice (24 percent). Model 4 shows the results hold when additional controls are added. Here, when Female_t , Male_{t+1} is set to 1 and all other variables are held constant at their mean and modal values there is a 40 percent chance the female Justice is interrupted, as compared to a 29 percent chance when Female_t , Male_{t+1} is set to 0.⁵

⁵ We also note that the absolute difference in ideology is a significant negative predictor of vocal pitch

Table 4: Male Justices Raise Their Vocal Pitch When Answering Female Lawyers

	<i>Dependent variable:</i>			
	Vocal Pitch _{t+1}		Interrupted _{t+1}	
	(1)	(2)	(3)	(4)
Fixed Effects				
Constant	−0.062*** (0.005)	−0.098*** (0.010)	−1.124*** (0.016)	−1.063*** (0.029)
Female _t , Male _{t+1}	0.149*** (0.015)	0.144*** (0.015)	0.111** (0.046)	0.181*** (0.047)
Minutes _t		0.001*** (0.0002)		0.001** (0.0004)
Ideology _{t+1}		−0.016*** (0.003)		−0.116*** (0.009)
Liberal Lower Court Decision		−0.015** (0.007)		0.054*** (0.016)
Ideology _{t+1} × Liberal Lower Court Decision		0.003 (0.003)		−0.009 (0.009)
Seniority _{t+1} -Seniority _t		0.002*** (0.0005)		−0.002 (0.001)
Chief Justice _{t+1}		−0.005 (0.014)		−0.649*** (0.042)
Sex Discrimination		−0.032 (0.028)		0.026 (0.070)
Random Effects				
Dyad	0.109*** (0.331)	0.108*** (0.329)	1.231*** (1.110)	1.246*** (1.116)
N ₁	134,391	130,451	139,837	135,757
N ₂	11,741	11,455	12,309	12,023
Log Likelihood	−185,936.400	−180,432.100	−84,352.080	−81,222.160
AIC	371,880.700	360,886.300	168,710.200	162,464.300

Note: We use the subscripts t and $t + 1$ to indicate the speaker order, meaning Justice_t is speaking to Justice_{t+1}. Female_t, Male_{t+1} indicates a female Justice is speaking to a male Justice. In Models 1 and 2, we estimate a multilevel logistic regression given the outcome is whether the female Justice was interrupted (Interrupted_t). In Models 3 and 4, we estimated a multilevel linear regression since the outcome is the vocal pitch (in standard deviations) of the responding Justice ($t + 1$). Unit of analysis is each utterance with a random intercept for each dyad. Seniority_t is the number of years Attorney_t has appeared in front of the Court. Seniority_{t+1} is the number of years Justice_{t+1} has served. Levels of significance are reported as follows: * $p < .1$; ** $p < .05$; *** $p < .01$. Standard errors are reported in parentheses.

We also estimated models using female lawyers. These can be found in Table 4. As before, the unit of analysis is an individual utterance with t denoting the order of the interaction. Our primary independent variable is whether a female attorney was addressing a male Justice ($Female_t, Male_{t+1}$). Model 1 shows that male Justices tend to lower their vocal pitch when addressing male attorneys, but when female attorneys are addressing male Justices they tend to raise their vocal pitch. Not only is this a statistically significant relationship, but the “Lawyer” panels of Figure 3 show the difference is substantial. More specifically, when female Justices are addressing male Justices the responding vocal pitch is a little over twice as high. Essentially the same result is found when additional controls are added to the model, including whether the speaker is addressing the Chief Justice, the seniority difference between the answering Justice and speaker, and the ideology of the answering Justice. In order to gain some leverage on the ideology of the argument being advanced, we also included a dummy variable capturing whether the lower court issued a liberal ruling. Even with these controls, there is a statistically significant ($p < .05$) difference between the interactions of female attorneys and male Justices, as compared to all other combinations.

We conduct similar analyses in Models 3 and 4, except instead of determining whether a male Justice is more likely to raise the tone of his voice, we consider whether a male Justice is more likely to interrupt a female attorney. Figure 3 shows the probability of an interruption increases around 2.10 percentage points when a female attorney addresses a male Justice ($p < .05$). Although slight, this result implies that interruptions are significantly more likely when female attorneys are addressing male Justices as opposed to when female attorneys are addressing female Justices or male attorneys are addressing either male or female Justices.

and interruptions in Models 3 and 4. Specifically, the more ideologically apart the two speakers are, the less likely the second speaker will raise his voice to the first speaker. Similarly, as ideological distance grows the likelihood of an interruption occurring *decreases*. However, upon further examination this result does not seem to be particularly strong. For example, when a simple t -test is used to compare Justices on the same side of the ideological spectrum (i.e., both with MQ scores greater than 0) versus those that are not, no significant difference is found between the two groups. This result holds regardless of whether one considers changes in vocal pitch or interruptions. Moreover, when Models 3 and 4 are re-estimated using this binary variable, the variable is not statistically significant ($p > .05$), suggesting that there is no significant relationship between ideological difference and aggressive turn-taking behavior.

As before, this result holds in Model 4 even when additional controls are included, suggesting that male Justices seem to behave differently towards women, even though this relationship is more pronounced when they interact with other Justices.

5.2 Female Justices Withdraw From Oral Arguments When Male Justices Raise Their Vocal Pitch

Of course, the differences in male non-verbal behavior toward their female colleagues is, perhaps, only relevant if it changes their behavior. If women Justices, highly professional and, often, accomplished debaters, are un-fazed by the change of the tone from the men on the Court, then the behavior of the men is of little direct consequences. As such, we look for evidence that women change their way of speaking after a man raises his vocal pitch. Fortunately, the Supreme Court affords a unique opportunity for exploring these questions. Specifically, we have at our disposal several dozen years of interactions between the same group of individuals—a requirement impossible to match in experimental settings or even in educational ones. Thus, we can examine whether a female Justice’s behavior changes in response to increases in male vocal pitch.

In Table 5, we find strong evidence of the latter. In these models, the dependent variable is whether the female Justice’s vocal pitch *decreased* after a male Justice raised his vocal pitch. To construct this variable, we subtracted the vocal pitch of a given question at time t from the average vocal pitch of subsequent questions at times $t + 1, t + 2, \dots, t + n$. This was done for each interaction in our dataset. For example, Ruth Bader Ginsburg addressed Antonin Scalia 172 times in our data set. Antonin Scalia raised his vocal pitch in 80 of these interactions. Prior to Justice Scalia’s interruption in *Rodriguez v. United States*, Ruth Bader Ginsburg’s vocal pitch was around 1.07 standard deviations *below* her baseline. Afterwards, her vocal pitch dropped to 2.71 standard deviations *below* her baseline, suggesting she was *less* activated in oral arguments after Justice Scalia’s interruption. We are interested in whether similar decreases generally occur when male Justices raise their vocal pitch and

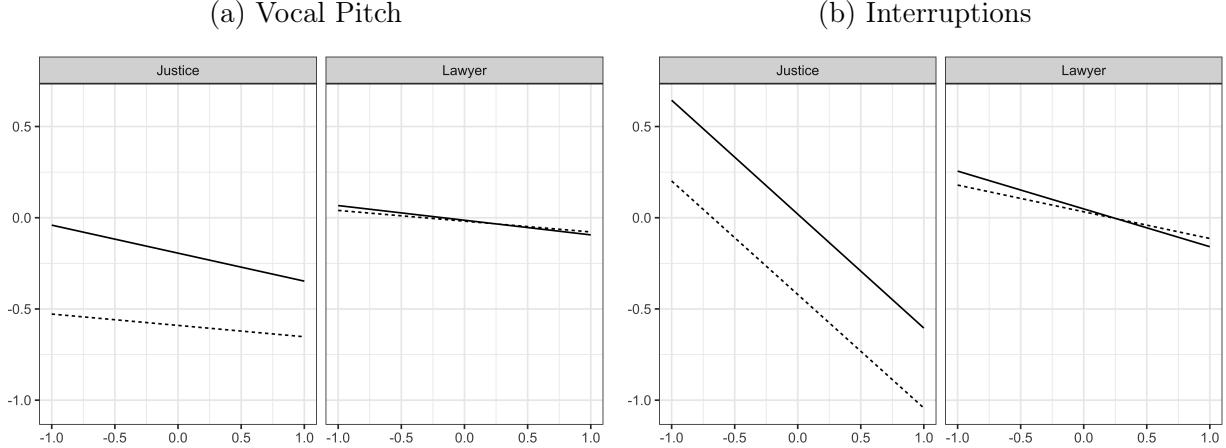
Table 5: When Male Justices Raise Their Vocal Pitch, Female Justices Become Emotionally Disengaged

	<i>Dependent variable:</i>			
	Remaining Vocal Pitch		Remaining Vocal Pitch	
	(1)	(2)	(3)	(4)
Fixed Effects				
Constant	−0.590*** (0.051)	−0.884*** (0.102)	−0.421*** (0.045)	−0.658*** (0.093)
Female _t , Male _{t+1}	0.396*** (0.109)	0.445*** (0.127)	0.440*** (0.100)	0.465*** (0.119)
Minutes _t		0.003** (0.001)		0.003*** (0.001)
Ideology _{t+1} − Ideology _t		0.022 (0.030)		0.009 (0.026)
Seniority _{t+1} -Seniority _t		−0.003 (0.004)		−0.001 (0.003)
Chief Justice _t		0.167 (0.145)		0.108 (0.125)
Chief Justice _{t+1}		0.188 (0.140)		0.168 (0.120)
Sex Discrimination		0.202 (0.170)		0.192 (0.161)
Vocal Pitch _{t+1}	−0.062*** (0.018)	−0.070*** (0.019)		
Female _t , Male _{t+1} × Vocal Pitch _{t+1}	−0.091** (0.041)	−0.087* (0.045)		
Interrupted _{t+1}			−0.622*** (0.046)	−0.574*** (0.051)
Female _t , Male _{t+1} × Interrupted _{t+1}			−0.002 (0.098)	−0.016 (0.109)
Random Effects				
Dyad	0.176*** (0.420)	0.197*** (0.444)	0.115*** (0.339)	0.133*** (0.364)
N ₁	6,909	5,992	7,316	6,355
N ₂	143	132	143	132
Log Likelihood	−12,603.210	−11,078.860	−13,218.150	−11,657.780
AIC	25,218.430	22,181.720	26,448.300	23,339.570

Note: Female_t, Male_{t+1} indicates a female Justice is speaking to a male Justice. In Models 1-4 we estimated a multilevel linear regression since the outcome is the difference between the questioning Justice's vocal pitch (in standard deviations) at time t and their average vocal pitch after ($t + 1, t + 2, \dots, t + n$) the responding Justice either raised his/her vocal pitch (Models 1 and 2) or interrupted them (Models 3 and 4). Positive values implies the questioning Justice's vocal pitch increased after the interaction. Negative values implies it decreased. Unit of analysis is each utterance with a random intercept for each dyad. Seniority_t and Seniority_{t+1} are the number of years Justice_t and Justice_{t+1} have served, respectively. Levels of significance are reported as follows: * $p < .1$; ** $p < .05$; *** $p < .01$. Standard errors are reported in parentheses.

interrupt female Justices and lawyers.

Figure 4: Predicted Values for Tables 5 and 6



Note: The “Justice” panels use the coefficients from the first and third models in Table 5. The “Lawyer” panels use the coefficients from the first and third models in Table 6. The solid and dash lines imply Female_t , Male_{t+1} was set to 0 and 1, respectively. Error bars represent 95 confidence intervals from 1,000 bootstrapped samples. The Supplemental Information contains more information regarding the bootstrapping procedure.

Model 1 demonstrates that female Justices tend to become disengaged (or de-activated) after male Justices raise their vocal pitch. To help, we plotted the predicted values in Figure 4 as the responding male Justice’s vocal pitch ranges from one standard deviation below his baseline (-1) to one standard deviation above his baseline (1). Here, when Female_t , Male_{t+1} is set to 1 and Vocal Pitch_{t+1} is set to 0 (or at the male Justice’s baseline), the predicted decrease in vocal pitch is around 0.19 standard deviations. When Vocal Pitch_{t+1} is set to 1 (or one standard deviation above the male Justice’s baseline), the predicted decrease nearly doubles to 0.35 standard deviations below the female Justice’s baseline. This result holds when additional controls are used in Model 2. Here, when all other variables are set to their mean and modal values a standard deviation increase in the male Justice’s vocal pitch corresponds with a 0.16 decrease in the female Justice’s vocal pitch for the remainder of oral arguments.

We find essentially the same results for interruptions. When female Justices are interrupted by male Justices (Female_t , $\text{Male}_{t+1} = 1$ and $\text{Interrupt}_{t+1} = 1$) their vocal pitch

decreases 0.60 standard deviations. When they are not interrupted, their vocal pitch actually increases 0.02 standard deviations, suggesting interruptions have a stronger affect on female Justice behavior as compared to subtle increases in vocal pitch. As before, these results hold when additional controls are included for Justice ideological and seniority differences as well as whether either Justice was Chief Justice at the time of the interaction. When these variables are held at their mean and modal values, a female Justice's vocal pitch decreases 0.67 standard deviations after being interrupted, whereas it only decreases 0.08 standard deviations if an interrupt does not occur. Again, this suggests that interruptions have a greater influence on female Justice activation as compared to changes in vocal pitch – something that is readily apparent when one compares the Justice slopes in Figure 4.

Table 6 finds similar results when female lawyers interact with male Justices, although the relationship seems to be less pronounced. For example, when Female_t , Male_{t+1} is set to 1 and the vocal pitch of the male Justice (Vocal Pitch_{t+1}) is set to 0, female lawyers decrease their vocal pitch 0.01 standard deviations for the remainder of oral arguments. When Vocal Pitch_{t+1} is set to 1, female lawyers further decrease their vocal pitch around 0.08 standard deviations. This is about half the difference we found when we did the same calculation for female Justices. With that said, the result still holds when additional controls are included. Here, when all additional variables are set to their mean and modal values a standard deviation increase in male Justice vocal pitch corresponds with a 0.08 standard deviation decrease in the vocal pitch of female lawyers.

Collectively, this suggests that when male Justices increase their vocal pitch that female lawyers become less engaged in oral arguments. Similar to female Justices, Models 3 and 4 show interruptions seem to have a more pronounced effect on the vocal pitch of female lawyers. Here, when female lawyers are interrupted by male Justices (Female_t , $\text{Male}_{t+1} = 1$ and $\text{Interrupt}_{t+1} = 1$) their vocal pitch *decreases* 0.16 standard deviations for the remainder of oral arguments. Conversely, when they are not interrupted, their vocal pitch *increases* 0.05 standard deviations. When the control variables found in Model 4 are held constant at

Table 6: When Male Justices Raise Their Vocal Pitch, Female Lawyers Become Emotionally Disengaged

	<i>Dependent variable:</i>			
	Remaining Vocal Pitch		Remaining Vocal Pitch	
	(1)	(2)	(3)	(4)
Fixed Effects				
Constant	−0.019*** (0.004)	0.012 (0.008)	0.033*** (0.004)	0.068*** (0.009)
Female _t , Male _{t+1}	0.005 (0.012)	0.006 (0.012)	0.016 (0.014)	0.018 (0.014)
Minutes _t		−0.002*** (0.000)		−0.002*** (0.000)
Ideology _{t+1}		−0.007*** (0.003)		−0.008*** (0.003)
Liberal Lower Court Decision		−0.001 (0.007)		−0.00000 (0.006)
Ideology _{t+1} × Liberal Lower Court Decision		0.001 (0.003)		−0.001 (0.003)
Seniority _{t+1} -Seniority _t		0.000 (0.000)		−0.000 (0.000)
Chief Justice _{t+1}		0.107*** (0.011)		0.098*** (0.010)
Sex Discrimination		0.020 (0.028)		0.009 (0.028)
Vocal Pitch _{t+1}	−0.059*** (0.003)	−0.058*** (0.004)		
Female _t , Male _{t+1} × Vocal Pitch _{t+1}	−0.021** (0.010)	−0.022** (0.010)		
Interrupt _{t+1}			−0.146*** (0.007)	−0.144*** (0.007)
Female _t , Male _{t+1} × Interrupt _{t+1}			−0.061*** (0.022)	−0.061*** (0.022)
Random Effects				
Dyad	0.013*** (0.113)	0.012*** (0.107)	0.012*** (0.110)	0.011*** (0.105)
N ₁	108,393	105,257	111,980	108,740
N ₂	11,113	10,833	11,223	10,943
Log Likelihood	−155,116.600	−150,674.400	−160,470.300	−155,882.600
AIC	310,245.300	301,374.700	320,952.500	311,791.100

Note: Female_t, Male_{t+1} indicates a female lawyer is speaking to a male Justice. Outcome is difference between the questioning lawyers's vocal pitch (in standard deviations) at time t and their average vocal pitch after ($t + 1, t + 2, \dots, t + n$) the responding Justice either raised his/her vocal pitch (Models 1-2) or interrupted them (Models 3-4). Unit of analysis is each utterance with a random intercept for each dyad. Seniority_t is the number of years Attorney_t has appeared in front of the Court. Seniority_{t+1} is the number of years Justice_{t+1} has served. Levels of significance are reported as follows: * $p < .1$; ** $p < .05$; *** $p < .01$.

their mean and modal values female lawyers decrease their vocal pitch around 0.21 standard deviations *more* when they are interrupted by male Justices as compared to when they are not. Again, this suggests that both vocal pitch and interruptions affect the non-verbal behavior of female lawyers, but the effect of the latter seems to be more pronounced.

Although these results are consistent with the results found in Table 5, we want to make sure and not overstate the lawyer results. More specifically, Figure 4 shows that there is a very slight difference between the predicted values when Female_t , Male_{t+1} is set to 1 (see solid line) versus when it is set to 0 (see dashed line), suggesting female lawyers are less responsive to the behavior of male Justices. Although we cannot say definitively why this disparity exists, we suspect it has to do with the voluntary nature of female Justice participation. Unlike female lawyers, female Justices are not required to interact with male Justices. Female lawyers on the other hand have a strong incentive to carry on, regardless of the demeanor of the Justices they are interacting with. With that said, the lawyer and Justice results are consistent, suggesting the disengagement that occurs when male Justices raise their vocal pitch cannot be attributed to a few idiosyncratic relationships.

According to Karpowitz and Mendelberg (2014), women often have to display more masculine traits in order to be successful in professional environments. Consequently, “once this tough-as-nails, superconfident, highly qualified woman achieves a leadership role, she may not be much influenced by what men around her do – because she already acts in traditionally masculine ways” (355). The results outlined above are consistent with this argument. Sandra Day O’Connor, Ruth Bader Ginsburg, Sonia Sotomayor, and Elana Kagan have all faced gender discrimination inside and outside of the courtroom, meaning they have had to develop incredibly thick skin over the course of their careers. Ultimately, this suggests the effect of gender on the Supreme Court is likely very subtle. An increase in vocal pitch is unlikely to cause female Justices to completely withdraw from oral arguments, but it may cause them to become less engaged. Our results provide evidence consistent with this claim and demonstrate female Justices and lawyers seem to become more “reticent” and “inhibited”

in response to aggressive non-verbal turn-taking. In the next section, we discuss the broader descriptive and substantive implications of our results.

6 Concluding Remarks

Whether it is small group deliberations (Karpowitz, Mendelberg and Shaker, 2012; Karpowitz and Mendelberg, 2014), legislative committees (Kathlene, 1994; Rosenthal, 2000), or oral arguments (Feldman and Gill, 2017), scholars have been interested in how men and women interact during deliberations. Out of necessity, however, this literature has focused almost exclusively on interactions that take place either in laboratories or outside of elite institutions, and even work that examines elite interactions has focused on verbal communications, which can often be circumscribed by institutional norms. Our work, which focuses on a rather unique political environment, considers whether men and women display different non-verbal expressions even in elite settings. In doing so, we not only gain a better understanding of gender dynamics, but also evaluate how implicit gender bias can have downstream consequences.

We find, consistent with studies in other contexts, that gender—and gender dynamics—play an important role in structuring how the Justices interact with one another. Specifically, we find that male Justices are more likely to raise their vocal pitch when addressing female Justices. We also find that, when male Justices answer female Justices, their vocal pitch tends to be higher than if they were answering other male Justices, suggesting that male Justices generally are “talking over” female Justices in an attempt to exert control of the conversation. We find similar patterns with regards to interactions between male Justices and female attorneys. Indeed, male Justices are more likely to raise their vocal pitch when both questioning and answering female lawyers, suggesting these gendered patterns are not idiosyncratic to a handful of individuals.

Feldman and Gill (2017) argue that an “an increasing number of interruptions may lead to diminished engagement” of the interrupted Justices (62). We find evidence that female

Justices and lawyers tend to become more disengaged after male Justices raise their tone of voice. This implies even subtle changes in vocal inflections can have important downstream consequences on subsequent participation by female Justices. If justices “possess unique and valuable information emanating from shared professional experiences” (Boyd, Epstein and Martin, 2010, 391) then such “diminished engagement” carries important substantive implications for how women’s interests are adjudicated on the Supreme Court.

Similarly, our results also carry considerable descriptive weight. If Sandra Day O’Connor, Ruth Bader Ginsburg, Sonia Sotomayor, and Elena Kagan are “talked over” by their male colleagues, then other – less accomplished – women likely face greater obstacles when deliberating in their professional environments, especially if those deliberations take place behind closed doors. “For most of our nation’s history, the concepts ‘woman’ and ‘lawyer’ were thought incompatible” (Ginsburg, 1978, 2). With the number of female judges doubling (Rhode, 2001), “woman” and “judge” are also becoming seemingly more compatible, but like other elite women our results demonstrate there is considerable progress that still needs to be made.

Oral arguments are the only time when Justices get to question one another, making them incredibly useful for learning about the case and the positions of the other Justices. If female Justices are systematically excluded from this process, then they are less able to introduce their own opinions into oral arguments. This not only puts them at an informational disadvantage, but it may also exclude them from coalitions that form after the oral arguments conclude. Over the course of their careers, many of the female Justices have been staunch advocates of women’s rights which means their exclusion not only affects their ability to influence Court decisions, but it may also affect the ability of women to receive a favorable hearing in front of the Court.

Unlike Congress, in which floor speeches are usually planned and scripted (and oriented toward a broader public audience), the interactions between the Justices at oral argument are unstructured, intimate, and, compared to Congressional floor exchanges, fairly

contemporaneous. In addition, and also unlike Congress, the Justices operate as co-equals, and, although seniority is important, no hierarchy dictates the nature of oral arguments proceedings; neither is there political or professional competition, which some scholarship (e.g., Kanthak and Woon, 2015) has shown deters female participation. This unstructured dialogue therefore provides a good—and fairly unique within the elite levels of American politics—setting in which to understand gender dynamics.

Unfortunately, during oral arguments male and female Justices are far from equal. When addressing female Justices male Justices raise the tone of their voice. Male Justices do the same when they answer female Justices. The same patterns are found when male Justices interact with female attorneys, suggesting there is likely an underlying gender dynamic governing how men and women engage one another during oral arguments. Female Justices face inherent disadvantages when they participate in oral arguments which is likely why they become less engaged after male Justices raise the tone of their voice. Further research is needed to fully understand this relationship, but our results demonstrate that regardless of the research scholars conduct, non-verbal cues should be actively considered – especially when considering gender dynamics in elite settings where important institutional norms prohibit the display of certain types of behavior, like interruptions. Until this is done, we may not fully appreciate the role gender plays in political deliberations, including those that take place on the Supreme Court.

References

- Anderson, Kristin J. and Campbell Leaper. 1998. “Meta-Analyses of Gender Effects on Conversational Interruption: Who, What, When, Where, and How.” *Sex Roles* 39(3-4):225–252. 4
- Baxter, Judith. 2014. *Double-Voicing at Work: Power, Gender and Linguistic Expertise*. Springer. 4

Beattie, Geoffrey W, Anne Cutler and Mark Pearson. 1982. "Why is Mrs. Thatcher Interrupted so Often?" *Nature* 300(23/30):744–747. 6

Beckham, B. and H. Aronson. 1978. "Selection of Jury Foremen as a Measure of the Social Status of Women." *Psychological Reports* 43(2):475–478. 8

Black, Ryan C., Sarah A. Treul, Timothy Johnson and Jerry Goldman. 2011. "Emotions, Oral Arguments, and Supreme Court Decision Making." *The Journal of Politics* 73(2):572–581.

1

Boersma, Paul. 1993. "Accurate Short-Term Analysis of the Fundamental Frequency and the Harmonics-to-Noise Ratio of a Sampled Sound." *Proceedings of the Institute of Phonetic Sciences* 17:97–110. 12

Boyd, Christina L ., Lee Epstein and Andrew D. Martin. 2010. "Untangling the Causal Effects of Sex on Judging." *American Journal of Political Science* 54(2):389–411. 8, 34

Burger, Warren E. 1975. "The Necessity for Civility." *Litigation* 1:8. 5

Calvo, Rafael A. and Sidney D'Mello. 2010. "Affect Detection: An Interdisciplinary Review of Models, Methods, and Their Applications." *IEEE Transactions on Affective Computing* 1(1):18–37.

Cutler, Anne and Mark Pearson. 1986. "On the Analysis of Prosodic Turn-Taking Cues." *Intonation in Discourse* pp. 139–155. 6

Duncan, Starkey and Donald W Fiske. 2015. *Face-to-face interaction: Research, methods, and theory*. Vol. 3 Routledge. 6

Ekman, Paul, Maureen O'Sullivan, Wallace V. Friesen and Klaus R. Scherer. 1991. "Invited Article: Face, Voice, and Body in Detecting Deceit." *Journal of Nonverbal Behavior* 15(2):125–135. 7

Epstein, Lee, Jack Knight and Andrew D. Martin. 2003. "The Norm of Prior Judicial Experience and Its Consequences for Career Diversity on the U.S. Supreme Court." *California Law Review* 91(4):903–965.

Feldman, Adam and Rebecca D. Gill. 2017. "Echoes from a Gendered Court: Examining the Justices' Interactions During Supreme Court Oral Arguments.". 5, 7, 8, 33

Fowler, Lucy. 2005. "Gender and Jury Deliberations: The Contributions of Social Science." *William & Mary Journal of Women and the Law* 12:1–48. 8

Gallafent, A. 2008. "The Sound of Leadership.". 6

Ginsburg, Ruth Bader. 1978. "Women at the Bar-A Generation of Change." *University of Puget Sound Law Review* 2:1. 34

Hancock, Adrienne B. and Benjamin A. Rubin. 2015. "Influence of Communication Partner's Gender on Language." *Journal of Language and Social Psychology* 34(1):46–64. 4

Johnson, Timothy R., Ryan C. Black and Justin Wedeking. 2009. "Pardon the Interruption: An Empirical Analysis of Supreme Court Justices' Behavior During Oral Arguments." *Loyola Law Review* 55:331–352. 5, 7

Kanthak, Kristin and Jonathan Woon. 2015. "Women don't run? Election aversion and candidate entry." *American Journal of Political Science* 59(3):595–612. 1, 35

Karpowitz, Christopher F. and Tali Mendelberg. 2014. *The Silent Sex: Gender, Deliberation, and Institutions*. Princeton University Press. 4, 32, 33

Karpowitz, Christopher F., Tali Mendelberg and Lee Shaker. 2012. "Gender Inequality in Deliberative Participation." *American Political Science Review* 106(03):533–547. 1, 4, 33

Kathlene, Lyn. 1994. "Power and Influence in State Legislative Policymaking: The Interaction of Gender and Position in Committee Hearing Debates." *American Political Science Review* 88(3):560–576. 4, 33

Kimmel, Christopher M, Patrick A Stewart and William D Schreckhise. 2012. "Of Closed Minds and Open Mouths: Indicators of Supreme Court Justice Votes During the 2009 and 2010 Sessions." 10(2). 5

Ko, Sei Jin, Melody S. Sadler and Adam D. Galinsky. 2014. "The Sound of Power Conveying and Detecting Hierarchical Rank Through Voice." *Psychological Science* 26:3–14. 6

Kollock, Peter, Philip Blumstein and Pepper Schwartz. 1985. "Sex and Power in Interaction: Conversational Privileges and Duties." *American Sociological Review* pp. 34–46. 4

Marder, Nancy S. 1987. "Gender Dynamics and Jury Deliberations." *The Yale Law Journal* 96(3):593–612. 8

Mauss, Iris B. and Michael D. Robinson. 2009. "Measures of Emotion: A Review." *Cognition and Emotion* 23(2):209–237. 6, 8

Nemeth, Charlan, Jeffrey Endicott and Joel Wachtler. 1976. "From the '50s to the '70s: Women in Jury Deliberations." *Sociometry* pp. 293–304. 8

Pomerantz, Eva M., Ellen Rydell Altermatt and Jill L. Saxon. 2002. "Making the Grade But Feeling Distressed: Gender Differences in Academic Performance and Internal Distress." *Journal of Educational Psychology* 94(2):396. 4

Posner, Jonathan, James A. Russell and Bradley S. Peterson. 2005. "The Circumplex Model of Affect: An Integrative Approach to Affective Neuroscience, Cognitive development, and Psychopathology." *Development and Psychopathology* 17:715–734. 7

Rhode, Deborah L. 2001. *The Unfinished Agenda: Women and the Legal Profession*. Chicago, IL: American Bar Association, Commission on Women in the Profession. 34

Rosenthal, Cindy Simon. 2000. "Gender Styles in State Legislative Committees: Raising Their Voices in Resolving Conflict." *Women & Politics* 21(2):21–45. 4, 33

Rosenthal, Cindy Simon, Jocelyn Jones and James A Rosenthal. 2003. "Gendered Discourse in the Political Behavior of Adolescents." *Political Research Quarterly* 56(1):97–104. 2

Spaeth, Harold J., Lee Epstein, Andrew D. Martin, Jeffrey A. Segal, Theodore J. Ruger and Sara C. Benesh. 2015. "Supreme Court Database, Version 2015, Release 01." <http://www.supremecourtdatabase.org> (Accessed on August 23, 2015). 17

Strodtbeck, Fred L. and Richard D. Mann. 1956. "Sex Role Differentiation in Jury Deliberations." *Sociometry* 19(1):3–11. 8

Swann, Joan and David Graddol. 1988. "Gender Inequalities in Classroom Talk." *English in Education* 22(1):48–65. 2

Thomas, Clarence. 1996. "Civility and Public Discourse." *New England Law Review* 31:515. 5

Thomas, Clarence. 1997. "A Return to Civility." *Tulsa Law Journal* 33:7. 5

Thomas, Clarence. 1998. "Civility-A Speech Delivered by Associate Justice Clarence Thomas to Students at Washington and Lee University School of Law-Lexington Virginia-Tuesday, March 10, 1998." *Washington and Lee Journal of Civil Rights and Social Justice* 4:1. 5

Toro, Isabella M. Garcia. 2015. "Gender Matters in Jury Duty: Male and Female Participation during the Deliberation Process." *Revista Juridica UPR* 84:61–78. 8

Zellers, Margaret. 2013. "Prosodic Variation for Topic Shift and other Functions in Local Contrasts in Conversation." *Phonetica* 69(4):231–253. 6

Zuraidah, Mohd Don and Gerry Knowles. 2006. "Prosody and Turn-Taking in Malay Broadcast Interviews." *Journal of Pragmatics* 38(4):490–512. 6

7 Supplemental Information

This Supplemental Information includes additional analyses for the following article:

Dietrich, Bryce J., Ryan D. Enos, and Maya Sen. “Gender Dynamics in Elite Political Contexts: Evidence from Supreme Court Oral Arguments.”

Description of Supreme Court Audio Data

Although the U.S. National Archives maintain the official audio recordings, for this paper we accessed the audio files in .mp3 format via the Oyez Project at the Chicago-Kent College of Law, which is a public repository that aims “to be a complete and authoritative source for all audio recorded in the Court since the installation of a recording system in October 1955.” To get additional information on the content of the files and to verify the identity of each speaker, we also obtained the transcript for each argument which can be found in the Oyez Project player.

As explained on Oyez⁶, we do not have audio for every case. For example, some of the audio from the 1970s and 1990s resides in the National Archives in tape format. Even though Oyez is actively transferring this to a digital format, they still have not completed this process. There was also a problem with the Court’s tape machine, especially during the early 1990s, which made the Justices talk slightly faster than normal. In order to synchronize these recordings with the transcripts, the Oyez team (especially Pat Ward) had to manually re-engineer the audio to make it “faithful to the speakers’ voices.”

Other problems arose in the 1980s. Unbeknownst to the Court, their tape manufacturer changed the formula they used to create the reel-to-reel tapes used for oral arguments. Instead of lasting 80 years, like typical reel-to-reel tapes, these new tapes only lasted a few years. Ultimately, they were plagued by what Oyez calls the “sticky shed syndrome” in which tape reels tended to stick together when stored. To solve this problem, the Oyez team had

⁶<https://www.oyez.org/about>

to bake the tapes in a slow oven for several hours, then mount the tapes and dub them to another real or digital storage device. Unfortunately, this was only partially successful and many of the recordings from the 1980s were lost

These and other problems described in the “About the audio” portion of the Oyez website means the majority of our dataset is from 1998-2014. As these are essentially random errors, we use the full dataset in all of our models. For example, the Justices were unaware that the manufacturer had changed the composition of the reel-to-reel tapes. Similarly, we are also confident the Justices did not change their behavior based on a recording error that was uncovered decades later. Moreover, our results are robust to excluding years where we have a number of missing cases (i.e., 1980s and early 1990s). Indeed, our results remain the same regardless of whether our models are estimated using data from 1981-2014 or 1998-2014.

Below we provide additional detail about the data collection process.⁷

Vocal Pitch Across Lawyers and Dyads

⁷Our data collection took place between June 1, 2015 and August 1, 2015. Sometime between September 6, 2015 and October 12, 2015 Oyez changed their website dramatically. The new version of the website contains more aggressive web scraping barriers. We provide links to a version of the website captured on July 1, 2015, even though we acknowledge this is not reflective of the current site. We chose this version of the website because it is reflective of the website we worked with in order to collect our data.

Table S1: Average vocal pitch and standard deviation for lawyers (male lawyers), 1982–2014.

Lawyer	Appearances	Pitch Mean	Pitch SD
Daniel A. Small	121	298.12	42.23
Ernest D. Preate, Jr.	104	258.94	69.65
Dale M. Cendali	144	239.78	22.25
Frank Cicero, Jr.	118	232.97	49.65
Jack A. Wheat	150	232.28	37.79
C. C. Harness, III	91	223.61	23.90
Eugene E. Stearns	92	223.20	20.91
William L. Williams	58	219.75	53.71
Abraham S. Alter	73	204.84	13.95
Paul D. Boas	102	204.56	27.12
Michael Gordon	109	199.72	12.38

Note: Measurements of vocal pitch are in Hertz (Hz). To calculate each we used all the audio from lawyers questions and non-questions. We restricted the table to lawyers who appeared at least 10 times. The table is sorted by the mean vocal pitch.

Table S2: Average vocal pitch and standard deviation for lawyers (female lawyers), 1982–2014.

Lawyer	Appearances	Pitch Mean	Pitch SD
Sheila L. Birnbaum	194	322.50	40.01
Deborah Jane Chuang	81	243.38	30.76
Vicki C. Jackson	48	237.45	21.25
Lisa S. Gochman	116	237.03	16.44
Mary E. Maguire	93	235.33	12.71
Paula K. Smith	95	234.75	18.94
Theresa R. Wilson	113	234.33	25.09
Donna R. Andrieu	176	234.19	24.57
Elizabeth S. Kertscher	114	234.19	14.17
Elizabeth A. Johnson	63	233.89	13.74

Note: Measurements of vocal pitch are in Hertz (Hz). To calculate each we used all the audio from lawyers questions and non-questions. We restricted the table to lawyers who appeared at least 10 times. The table is sorted by the mean vocal pitch.

Table S3: Average vocal pitch and standard deviation for Justice-Justice dyads, 1982–2014.

First Justice	Second Justice	Appearances	Pitch Mean	Pitch SD
David H. Souter	Sandra Day O'Connor	110	191.41	45.23
Elena Kagan	Sonia Sotomayor	54	188.19	33.70
Sonia Sotomayor	Elena Kagan	88	188.05	42.62
Sandra Day O'Connor	Ruth Bader Ginsburg	70	185.08	40.17
Sonia Sotomayor	Ruth Bader Ginsburg	78	178.12	36.33
Samuel Alito	Sonia Sotomayor	32	177.59	42.67
Ruth Bader Ginsburg	Sandra Day O'Connor	60	176.86	39.27
Antonin Scalia	Elena Kagan	116	176.22	34.85
Ruth Bader Ginsburg	Elena Kagan	54	176.05	26.83
Anthony Kennedy	Sonia Sotomayor	16	175.24	36.33

Note: Measurements of vocal pitch are in Hertz (Hz). To calculate each we used all the audio from questions and non-questions. We restricted the table to dyads which appeared at least 10 times. The table is sorted by the mean vocal pitch. Women tend to talk at a higher vocal pitch. In future drafts, we will use male-male/female-female dyads.

Table S4: Average vocal pitch and standard deviation for Justice-lawyer dyads, 1982–2014.

First Speaker	Second Speaker	Appearances	Pitch Mean	Pitch SD
Stephen G. Breyer	Daniel A. Small	20	259.81	62.15
Stephen G. Breyer	Debra J. Dixon	18	256.54	58.74
Stephen G. Breyer	Frank Cicero, Jr.	14	240.04	34.67
Sandra Day O'Connor	Kathryn Grill Graeff	12	238.46	78.63
Sandra Day O'Connor	C. C. Harness, III	18	235.41	23.45
Sandra Day O'Connor	Ernest D. Preate, Jr.	46	235.17	33.23
Sandra Day O'Connor	Dale M. Cendali	12	233.50	27.50
Ruth Bader Ginsburg	Sheila L. Birnbaum	26	230.61	84.50
David H. Souter	Sheila L. Birnbaum	18	228.35	101.43
John Paul Stevens	Ernest D. Preate, Jr.	24	226.78	78.81

Note: Measurements of vocal pitch are in Hertz (Hz). To calculate each we used all the audio from questions and non-questions. We restricted the table to dyads which appeared at least 10 times. The table is sorted by the mean vocal pitch. Women tend to talk at a higher vocal pitch. In future drafts, we will use male-male/female-female dyads.