Online Appendix for

Gender identity and mode-switching behavior: Evidence from the human voice

Yosh Halberstam

Correspondence to: <u>yosh.halberstam@gmail.com</u>

This PDF file includes:

Supplementary Text Figures S1 to S5 Tables S1 to S3

Contents

3
3
1
5
5
5
7
7
3
l
2
1
1
5
5
5
7
3
)

Section 1. Data Collection Process for Worker Samples

Section 1.1. Main Dataset: Lawyers from Vault 100 Law Firms

The Vault 100 is an annual ranking of the most prestigious law firms in the U.S. I consider law firms that were listed in the Vault 100 rankings at least once between 2016 and 2018. Based on information from a pilot sample, I dropped firms that had a live receptionist 24/7 and firms where less than 10 percent of voicemail greetings were self-recorded by lawyers. This left me with 84 law firms. Table S1 provides a list of these law firms.

From this list of 84 top private law firms, I used web scraping to collect the directory of lawyers' names and phone numbers. From May 2017 to January 2018, I collected 57,064 distinct phone numbers and used Voicent, an automated phone-calling software, to call each number during non-working hours (typically 2-5 AM EST), and to record each call that was successfully connected. I recorded the first 10 seconds of each successful call and extracted the first 3 seconds of each recording to minimize the likelihood of capturing silence or machine generated audio, such as generic instructions for leaving a message.

To detect poor quality clips, I decomposed each 3-second clip into 225 subintervals. Specifically, each clip is represented as a time interval [0, 3] in seconds, and each subinterval in that clip is defined as $\left[\frac{3k}{225}, \frac{3(k+1)}{225}\right]$ for $k \in \{0, ..., 224\}$. I used Praat's "To Intensity" function on the upper bound of each subinterval. This function, which estimates the amplitude at a given point, returns the value -300 dB if the clip is silent. For each clip, if more than 30% of the 225 sample points had an intensity of -300 dB, I eliminated the clip from my sample. A substantial number of clips were of poor quality, for example, due to unnatural acceleration or fragmentation of sounds in the clip. These issues often resulted in a high proportion of silence in the clip.

After eliminating unsuccessful recordings and poor quality clips, I extracted 39,962 lawyers' voicemail recordings from the 57,064 phone numbers. These comprise the main dataset of lawyers.

Section 1.2. Verified Female Lawyers from the Main Dataset

To determine which clips from the main dataset were self-recorded by a female lawyer, I listened to all clips from lawyers who were classified as female by ALM. If a clip was entirely recorded in first person by a human speaker, I classified the clip as self-recorded. This is in contrast to (1) assistant-recorded clips, which are in third-person, (2) ambiguous clips, which are human-recorded but do not contain first-person or third-person pronouns, (3) machine-recorded clips, which are clearly distinct from human voices, and (4) combinations of human-recorded and machine-recorded clips. I further eliminated poor-quality clips, for instance, clips that contained static, live answers, dialing noises, or unnatural acceleration of sounds. Through this process, I identified 6,618 self-recorded clips; however, 210 of them sounded male, and further information from their webpages, such as profile pictures and gender pronouns, confirmed this for 209 of them. I was unable to find gender information on one lawyer and omitted the clip from analysis.

Given these verification procedures, I decided to use the sample of 6,408 clips. The finite mixture models failed to converge on a solution for 9 of these clips, leaving me with a final sample of 6,399 clips that were verified as self-recorded by a female lawyer.

Section 1.3. Auxiliary Dataset: Law Firm Assistants

To find recordings of executive assistants of female lawyers, I listened to all clips from lawyers who were classified as female by ALM. To maximize the chances of finding recordings of female assistants of male lawyers, I listened to all clips which fulfilled all the following criteria:

1. The lawyers associated with these clips were classified as male by ALM.

2. The clips had a mean frequency above 150 Hz, well above the unitary male frequency mode of 100 Hz.

3. The clips had a probability of being self-recorded of at least 0.25, as determined by my machine learning classification (see Section 3 for details on the machine learning classification process).

I classified a clip as assistant-recorded if the clip contained third-person pronouns and was not recorded by a human. I eliminated clips that were recorded by non-executive assistants (i.e., the assistants who spoke generally on behalf of the firm rather than for a specific lawyer), clips that were fully or partially machine-recorded, and poor-quality clips. Through this process, I found 237 clips that were fully recorded by female executive assistants on behalf of female lawyers, and 412 clips that were fully recorded by female executive assistants on behalf of male lawyers.

Section 1.4. Auxiliary Dataset: Lawyers Who Switched Firms

I define "switchers" as lawyers who had switched firms since the initial collection of their voicemail greeting. As noted in Section 1.1, the period of initial collection was from May 2017 to January 2018. To find the switchers, about two years after the period of initial collection, I recruited MTurk workers to check the webpages of lawyers whose voicemails had a probability

of being self-recorded of at least 0.5, as determined by my machine learning classification (see Section 3 for details on the machine learning classification process). The MTurk workers were selected on a first-come-first-served basis from a pool of US-based MTurk workers who had a HIT approval rate of 99% or more and who had over 10,000 approved HITs.

I determined whether a lawyer had switched firms by clicking on their original profile page URL. If the URL linked to the lawyer's profile page, I determined that the lawyer had not switched firms. If the URL did not link to the lawyer's profile page, for instance by redirecting to a general directory or an error page, I determined that the lawyer had switched firms. See Figure S1 for the survey format.

I then recruited MTurk workers to find the switchers' new firms, profile page URLs, and personal office phone numbers. I provided MTurk workers with the switchers' names and old firms, and instructed respondents not to use sites other than Google, LinkedIn, and the lawyer's new firm website when providing information to avoid third-party websites of unknown reliability. See Figure S2 for details on the survey format and instructions.

As in Section 1.1, I used Voicent to collect voicemails from the new phone numbers. For each lawyer, I checked if both the old and new recordings were self-recorded by the same lawyer following the procedure in Section 1.2. In total, I found 627 lawyers with self-recorded voicemails at both their old and new firms. Of these 627 lawyers, 198 lawyers were female, as further verified using the procedure described above.

Section 1.5. Auxiliary Dataset: Promoted Lawyers

About two years after the period of initial data collection, I recruited MTurk respondents to check the job title of a subsample of lawyers from the lawyers' webpages. The workers were selected on a first-come-first-served basis from a pool of US-based MTurk workers who had a

HIT approval rate of 99% or more and who had over 10,000 approved HITs. I selected lawyers who fulfilled all the following criteria:

1. Confirmed to have stayed at their original firm.

2. Associates as of May 2017 – January 2018.

3. Had voicemails with a probability of being self-recorded of at least 0.5 (see Section 3 for details on the machine learning classification process).

I provided MTurk workers with the lawyers' website URLs, and instructed them to classify the lawyers' job title based on their personal webpage profiles. See Figure S3 for the survey format. This process elicited promotions data for 1,925 female lawyers whose clips were verified as self-recorded by a female lawyer using the procedure described earlier, and who were Associates as of May 2017 – January 2018. Of these 1,925 lawyers, 196 were promoted to Partner and 137 were promoted to Counsel/Other within the same firm as of January 2020.

Section 1.6. Auxiliary Dataset: Supreme Court Lawyers

From www.oyez.org, I collected data from 129 oral arguments made by female advocates at the U.S. Supreme Court between 1985 and 2005. From the recording of each argument, I took three voice samples and extracted the first 3 seconds of each sample. The samples are from the opening sentence, closing sentence, and one sentence taken from the middle of the argument (approximately minute 15).

Section 1.7. Auxiliary Dataset: RE/MAX Real Estate Agents

I collected the name, franchise, directory URL, and office number of 1,694 RE/MAX residential real estate agents through web scraping. These agents constituted all agents in Chicago, Dallas, Houston, and Phoenix in the fields of first-time buyers, luxury properties, condominiums, residential acreages, and rentals that were listed on <u>www.remax.com</u>. I also

collected the name, franchise, directory URL, and office number of 2,013 RE/MAX commercial real estate agents listed on <u>http://www.remaxcommercial.com/Roster/Agents</u> through web scraping.

I used Voicent to collect voicemails from these phone numbers, and manually classified these clips as self-recorded by the agent by listening to them. I successfully collected 539 selfrecorded voicemails from the residential real estate agents, and 527 self-recorded voicemails from the commercial real estate agents.

I then recruited MTurk respondents to check the gender of agents who had self-recorded voicemails. They were selected on a first-come-first-served basis from a pool of US-based MTurk workers who had a HIT approval rate of 99% or more and who had over 10,000 approved HITs. I provided MTurk respondents with the agents' names and directory URLs, and instructed them to classify the agents' gender based on their profile pictures. See Figure S4 for the survey format. Of the 539 residential real estate agents, 337 were classified as female. Of the 527 commercial real estate agents, 159 were classified as female.

Section 2. Estimating Frequencies from Audio Data

I used <u>Praat</u>, an open-source program to extract frequency data from each clip (Boersma, 1993). Praat chooses the frequency candidate associated with the highest *local* strength subject to various thresholds and the global path finder, a system that penalizes frequency variation across adjacent frames. This way, background noise and nonhuman sounds are less likely to confound estimates.

Specifically, to extract frequency data from each clip, I used the following baseline parameters for the function "To Pitch (ac)":

Pitch Floor: 50 Hz;
Pitch Ceiling: 400 Hz;
Time Step: 5 milliseconds;
Window: Hanning;
Silence Threshold: 0.03;
Voicing Threshold: 0.45;
Octave Cost: 0.01;
Octave-Jump Cost: 0.35;
Voiced/Unvoiced Cost: 0.14.

These parameters are defined by Praat as follows (Boersma, 1993; Boersma, 2001):

• *Pitch Floor:* The minimum frequency that will be considered for estimation. A pitch floor of 50 Hz is well below the range of human voice frequencies produced by the natural voice register. The pitch floor also determines the length of the analysis window, which is 3/*Pitch Floor* seconds long. With a pitch floor of 50 Hz, the analysis window is 0.06 seconds long.

• *Pitch Ceiling*: The maximum frequency that will be considered for estimation. A pitch ceiling of 400 Hz is well above the range of human voice frequencies produced by the natural voice register.

• *Time Step*: The interval between frequency estimates. The points in the clip where the first and the last frequency estimates are taken depend on the length of the analysis window, which in turn depends on the pitch floor. For a clip represented as a time interval [0, 3] in seconds, a pitch floor of 50 Hz and time step of 5 milliseconds mean that Praat produces one frequency estimate per 5 milliseconds in the subinterval

[0.03, 2.97], for a total of 589 frequency estimates including the endpoints of the subinterval.

I used the default Praat settings for the following:

• *Window*: I use the default Hanning window

• *Silence Threshold*: For each frame, if the local absolute amplitude peak is less than approximately Silence Threshold times the global absolute amplitude peak, the frame will be classified as "voiceless" (the frequency estimate will be a missing value). I use the default 0.03 as the silence threshold.

• *Voicing Threshold*: For each frame, if the strengths of all frequency candidates in the frame are below Voicing Threshold, the frame will be classified as "voiceless" (the frequency estimate will be a missing value). I use the default 0.45 as the voicing threshold.

• Octave Cost: This parameter determines how much higher-frequency candidates are favored relative to lower-frequency candidates. It is necessary to force Praat to choose a frequency candidate in the case of a perfectly periodic signal, where all autocorrelation peaks have equal values. I use the default 0.01 per octave as the octave cost.

• Octave-Jump Cost: This parameter determines the extent to which rapid pitch changes between adjacent frames are disfavored. In conjunction with Voiced/Unvoiced Cost, this is a global path finder parameter that affects estimates across rather than only within frames. I use the default 0.35 as the octave-jump cost.

• *Voiced/Unvoiced Cost*: This parameter determines the extent to which rapid transitions between voiced and voiceless frames are disfavored. In conjunction

with Octave-Jump Cost, this is a global path finder parameter that affects estimates across rather than only within frames. I use the default 0.14 as the voiced/unvoiced cost.

With these parameters, each clip comprises 589 frames. For each frame, the best frequency candidates were determined using the function "Get value in frame". Through this process, I obtained 589 frequency estimates for each clip, though many of these estimates are voiceless (i.e. missing values), and the number of such missing values differs across clips. See Table S3 for robustness checks on the values of the parameters used.

Section 3. Machine Learning Classification Process

I define a self-recorded clip as a clip which contains only a lawyer's voice, as opposed to an automated voice, an assistant's voice, or a combination of voices. Due to the large number of clips in the main dataset of 39,962 lawyers, I used machine learning to predict the probability that a clip is self-recorded for the main dataset. I used text information, acoustic information, and demographic information as predictors.

I obtained text information for each clip by transcribing the clips with IBM Watson Speech Recognition API. I decomposed the transcribed text into individual words, and selected the 50 most frequent words and the number of words per clip as predictors. I further manually selected 12 frequently occurring phrase patterns as predictors.

I obtained acoustic information for each clip using Praat and selected eight acoustic variables as predictors. I also obtained the demographic information of each lawyer from their websites. This information included: job title, practice area, law school, undergraduate school, any graduate degrees, graduation year for each degree earned, academic honors earned, and gender. Gender was assessed by scrutinizing lawyers' first names, subjectively classifying photos, searching for gendered pronouns in the lawyers' biographical descriptions, and cross-

referencing the recorded greeting. Generally, these gender indicators perfectly corroborated one another.

I set aside about 10% of the sample (the "ML sample") for training, testing, and validation. I manually classified all clips in the ML sample as self-recorded or otherwise following the procedure above. I randomly selected 80% of the ML sample for training and validation (the "training and validation sample") and the remaining 20% for testing ("the testing sample"). To address any possible overfitting issues, I used 5-fold cross validation to tune the model with the objective of minimizing logarithmic loss. I applied this procedure to four machine learning models: random forest, support vector machine, *k*-nearest neighbors, and XGBoost. XGBoost significantly outperformed the other three models with a predictive accuracy of 93.52%. With a probability threshold of 0.95, the XGBoost model had a predictive accuracy of 99.12%. I thus used XGBoost as my final machine learning model.

The most important non-acoustic variable was the number of words spoken by the lawyer in the greeting (coming in at sixth place). Gender was the most important, and indeed the only, demographic variable identified in the top-fifteen most predictive characteristics (coming in at eighth place).

Section 4. Survey Design for Relative Perceptions

To check if bimodality is correlated with perceptions of lawyers' characteristics, I surveyed 200 MTurk workers and asked them to listen to the same paired clips. After listening to each pair, workers were asked to rate the speakers on a relative scale on competitiveness, dominance, risk-taking attitude, seniority, and trustworthiness on a seven-point Likert scale. The survey instructions and format are shown in Figure S5. For each question, the position of the bimodal clip was randomized as either Recording 1 or Recording 2, and across questions for each HIT, the bimodal clip appeared 5 times as Recording 1 and 5 time as Recording 2. The order of characteristics on the Likert scale was also randomized for each question. These specifications, as well as any information about bimodality, were not made known to the workers.

Respondents were selected on a first-come-first-served basis from a pool of US-based MTurk workers who had a HIT approval rate of 99% or more and who had over 10,000 approved HITs. Only one HIT per respondent was considered; if the respondent completed more than one HIT, all additional HITs were rejected. I created male-only and female-only versions of the survey to achieve balance in respondent sex. In all, I had 100 male respondents and 100 female respondents across 200 HITs, with respondents self-reporting their sex.

Instructions:

There are 50 lawyers in this HIT.
 For each lawyer, click on the provided URL.
 If the URL opens the lawyer's personal webpage, click 'yes'. If the URL returns an error page or any other webpage, click 'no'.
 If you have accidentally checked the wrong box, click the other box to give the correct answer.

Please answer all questions to receive compensation.

Lawyer	Name	Url	Does the url lead to the lawyer's personal webpage?
1.	\${lawyer_name1}	search \${lawyer_name1}	Yes O No O
2.	\${lawyer_name2}	search \${lawyer_name2}	Yes O No O
3.	\${lawyer_name3}	search \${lawyer_name3}	Yes O No O

Figure S1. MTurk HIT for finding lawyers who switched firms

	Name	Old Firm Name	Click link to search:	New Personal Office Phone Number	New Personal Firm Page Url	New Firm Name	Comments
	Example	Case 1 & 2a:	All information found	123-456-7890	abcfirm.com/person/example1	ABC Firm	no comments
	Example	Case 2b:	Personal firm page not found	NA	linkedin.com/person/example2b	XYZ Firm	no comments
1.	\${lawyer_name1}	\${old_firm1}	search \${lawyer_name1}				
2.	\${lawyer_name2}	\${old_firm2}	search \${lawyer_name2}				
3.	\${lawyer_name3}	\${old_firm3}	search \${lawyer_name3}				

Figure S2. MTurk HIT for finding new office numbers

Instructions:

- There are 50 lawyers in this HIT.
 For each lawyer, click on the provided URL.
 Please classify the lawyer's gender and job title based on their profile.

Note: If a link does not lead to a lawyer's profile, please note this in the comments box, and choose 'uncertain' and 'counsel/other' respectively for the multiple choice questions.

Please answer all questions to receive compensation.

Lawyer	Url	What is the lawyer's gender?			What is the lawyer's job title?			Comments
1.	Lawyer Profile 1	\bigcirc Female	\bigcirc Male	◯ Uncertain	⊖ Associate	⊖ Partner	O Counsel/Other	
2.	Lawyer Profile 2	\bigcirc Female	\bigcirc Male	◯ Uncertain	⊖ Associate	⊖ Partner	O Counsel/Other	
3.	Lawyer Profile 3	⊖ Female	\bigcirc Male	◯ Uncertain	OAssociate	◯ Partner	○ Counsel/Other	

Figure S3. MTurk HIT for finding promoted lawyers

Instructions:

- There are 50 real estate agents in this HIT.
 For each agent, click on the provided URL.
 Please classify the agent's gender based on their website photo. Please make sure that you find the correct agent, and that you classify their gender based on their photo, not their name.

Note: If a link does not lead to an agent's profile, please note this in the comments box and choose 'uncertain' for the multiple choice question.

Please answer all questions to receive compensation.

Agent	Url	What is the agent's gender?	Comments
\${agent_1}	Agent Profile 1	○ Female ○ Male ○ Uncertain	
\${agent_2}	Agent Profile 2	○ Female ○ Male ○ Uncertain	
\${agent_3}	Agent Profile 3	◯ Female ◯ Male ◯ Uncertain	

Figure S4. MTurk HIT for real estate agents

	Instructions: Each audio clip in t For each attribute I For example, if the	his survey is a self-reco below, provide your relat worker in Recording 2 s	rded voicemail greeting of a p ive impression of the workers ounds more dominant than in	rofessional worker. I Recording 1, then select	a rating right of center.	
	lf, instead, in Reco Pair 1 of 10:	rding 1 the worker sound	ls significantly more dominan	t, then choose the leftmos	st option.	
	Recording 1:				Recording	2:
▶ 0:03 / 0	0:03 — 🔊	:			▶ 0:03 / 0:03	→ • :
			Risk-Taking			
0	0	0	0	0	0	0
			Competitiveness			
0	0	0	0	0	0	0
			Dominance			
0	0	0	0	0	0	0
			Trustworthiness			
0	0	0	0	0	0	0
			Seniority			
0	0	0		0	0	0

Figure S5. MTurk HIT for relative attribute ratings

Table S1. List of law firms

Akin Gump Strauss Hauer & Feld LLP Allen & Overy LLP Alston & Bird LLP Arent Fox LLP Arnold & Porter Kaye Scholer LLP BakerHostetler Baker McKenzie Ballard Spahr LLP Blank Rome LLP Boies Schiller Flexner LLP Bracewell LLP Bryan Cave Leighton Paisner LLP Cadwalader, Wickersham & Taft LLP Cahill Gordon & Reindel LLP Cleary Gottlieb Steen & Hamilton LLP Clifford Chance US LLP Covington & Burling LLP Crowell & Moring LLP **DLA Piper** Davis Polk & Wardwell LLP Davis Wright Tremaine LLP Debevoise & Plimpton LLP Dechert LLP Dentons Dorsey & Whitney LLP Faegre Drinker Biddle & Reath LLP Fenwick & West LLP Finnegan, Henderson, Farabow, Garrett & Dunner, LLP Fish & Richardson P.C. Foley & Lardner LLP Fox Rothschild LLP Gibson, Dunn & Crutcher LLP Goodwin Procter LLP Haynes and Boone, LLP Holland & Knight LLP Hunton & Williams LLP Irell & Manella LLP Jenner & Block LLP Jones Day K&L Gates LLP Kasowitz Benson Torres LLP Katten Muchin Rosenman LLP

Keker, Van Nest & Peters LLP Kilpatrick Townsend & Stockton LLP King & Spalding LLP Kirkland & Ellis Kramer Levin Naftalis & Frankel LLP Latham & Watkins LLP Locke Lord LLP Manatt, Phelps & Phillips, LLP Mayer Brown LLP McDermott Will & Emery Milbank LLP Morgan, Lewis & Bockius LLP Morrison & Foerster LLP Nixon Peabody LLP Norton Rose Fulbright LLP O'Melveny & Myers LLP Patterson Belknap Webb & Tyler LLP Paul Hastings LLP Paul, Weiss, Rifkind, Wharton & Garrison LLP Pepper Hamilton LLP Perkins Coie LLP Pillsbury Winthrop Shaw Pittman LLP Proskauer Rose LLP Ropes & Gray LLP Schulte Roth & Zabel LLP Seyfarth Shaw LLP Shearman & Sterling Sheppard Mullin Sidley Austin LLP Simpson Thacher & Bartlett LLP Skadden, Arps, Slate, Meagher & Flom LLP and Affiliates Squire Patton Boggs Steptoe & Johnson LLP Susman Godfrey LLP Troutman Pepper Hamilton Sanders LLP Venable LLP Vinson & Elkins LLP Weil, Gotshal & Manges LLP White & Case LLP Williams & Connolly LLP Wilson Sonsini Goodrich & Rosati Winston & Strawn LLP

	Group	Mean	Standard Deviation	Group	Mean	Standard Deviation
Total words	1	10.563	1.902	2	10.410	2.075
First word <i>hi</i> or <i>hello</i>	1	0.505	0.501	2	0.510	0.500
Includes reached	1	0.063	0.244	2	0.062	0.241
Includes at	1	0.042	0.201	2	0.038	0.192
Includes you've	1	0.042	0.201	2	0.044	0.204
Includes hi	1	0.035	0.184	2	0.036	0.186
Includes I'm	1	0.029	0.168	2	0.032	0.175
Unique words within group	1	0.248	0.432	2	0.216	0.411

 Table S2. Verbal content comparison

Notes: The table uses a random sample of 1,000 clips from the main sample of verified self-recorded greetings by female lawyers described earlier. Among these 3-second clips, 380 are classified as Group 1 and 620 as Group 2. The verbal content from each clip was extracted using an automated transcription program. Each row, except for the last, represents summary statistics for each group of clips. *Hi* or *hello* are the most frequently used first word in the greetings of each group. The subsequent five words, *reached*, *at*, *you've*, *hi*, and *I'm*, are the most commonly used words in the greetings of each group. Unique words refers to the share of words that appear only once within a group.

Frequency Estimates	Group	95% C.I. (Share)	95% C.I. (Location)	Group	95% C.I. (Share)	95% C.I. (Location)	Obs.
Residualized Baseline	1	(0.347, 0.380)	(-48.243, -46.326)	2	(0.620, 0.653)	(25.943, 28.044)	4,682
No Pathfinder	1	(0.611, 0.674)	(80.938, 83.021)	2	(0.326, 0.389)	(158.863, 164.568)	999
Stronger Pathfinder	1	(0.313, 0.378)	(83.361, 87.649)	2	(0.622, 0.687)	(160.376, 164.321)	998
Double OctaveCost	1	(0.309, 0.374)	(83.709, 87.763)	2	(0.626, 0.691)	(160.617, 164.596)	997
Gaussian Window	1	(0.330, 0.395)	(84.465, 88.315)	2	(0.605, 0.670)	(160.947, 164.894)	998
60 Hz Floor	1	(0.340, 0.408)	(84.672, 87.711)	2	(0.592, 0.660)	(157.858, 162.423)	997
40 Hz Floor	1	(0.297, 0.362)	(84.189, 88.385)	2	(0.638, 0.703)	(161.029, 164.884)	988

Table S3. Frequency	estimation	robustness	checks
---------------------	------------	------------	--------

Notes: The first row shows results using the residuals from regressing the location of the low mode (estimates from baseline 5-comp. FMM) on years of experience, firm, title, and litigator fixed effects. All subsequent rows use the baseline FMM. Rows 2-7 use the same random subset of 1,000 recordings described earlier. The second row switches off the pathfinder feature in Praat, which penalizes sharp changes in frequency. The location estimates are similar, but the distribution of groups is significantly different. The third row uses pathfinder parameters with increased values, which is the opposite of switching off the pathfinder. The fourth row uses double the *OctaveCost* as the baseline specification, thereby favoring higher-frequency candidates. The fifth row uses a Gaussian window to smooth the data rather than the default Hanning window. The sixth row uses a higher pitch floor, which implies a narrower analysis window of 50 milliseconds, and the seventh row uses a lower pitch floor, which implies a wider analysis window of 75 milliseconds.