

Appendix for the paper: "I would hire you in a minute: Thin slices of nonverbal behavior in job interviews"

Laurent Son Nguyen
 Idiap Research Institute
 École Polytechnique Fédérale de Lausanne
 Switzerland
 lnguyen@idiap.ch

Daniel Gatica-Perez
 Idiap Research Institute
 École Polytechnique Fédérale de Lausanne
 Switzerland
 gatica@idiap.ch

Appendix I: List of extracted nonverbal cues

In this appendix, we list all the extracted behavioral features used for the study. All features were extracted for all slices (and thin slice cases - question-only, answer-only, and whole-slice), as well as for the full interview. With the exception of silence/speech overlap, which refer to both protagonists jointly, all behavioral features were extracted for both the applicant and the interviewer. For details about the method used to extract these cues, please refer to the main paper. Table 1 and 2 display the list of audio features and visual features, respectively.

Table 1: List of extracted audio features.

<i>Silence/speech overlap</i>	
Number of silent events	
Time silent	
Number of overlapping speech events	
Time of overlapping speech	
<i>Speaking turns</i>	
Speaking time	
Number of turns	
Average turn duration	
Maximum turn duration	
Number of short utterances	
Short utterances time	
Number of pauses	
Number of audio backchannel	
Audio backchannel time	
<i>Prosody</i>	
Mean pitch	
Pitch STD	
Median pitch	
Pitch lower quartile	
Pitch upper quartile	
Maximum pitch	
Mean energy	
Energy STD	
Median energy	
Energy lower quartile	
Energy upper quartile	
Maximum energy	
Mean voiced rate	
Voiced rate STD	
Median voiced rate	
Voiced rate lower quartile	
Voiced rate upper quartile	
Maximum voiced rate	

Table 2: List of extracted visual features.

<i>Head visual motion</i>	
Mean head horizontal visual motion	
Horizontal head visual motion STD	
Median horizontal head visual motion	
Maximum horizontal head visual motion	
Mean vertical head visual motion	
Vertical head visual motion STD	
Median vertical head visual motion	
Maximum vertical head visual motion	
Mean magnitude of head visual motion	
Magnitude of head visual motion STD	
Median magnitude of head visual motion	
Maximum magnitude of head visual motion	
<i>Nodding</i>	
Number of nods	
Nodding time	
Number of visual backchannel events	
Visual backchanneling time	
Number of nodspeak events	
Nodspeak time	
<i>Overall motion (WMEI)</i>	
WMEI horizontal center of mass	
WMEI vertical center of mass	
WMEI mean	
WMEI median	
WMEI standard deviation	
WMEI lower quartile	
WMEI upper quartile	
WMEI maximum	
WMEI ratio of non-zero pixels	

Appendix II: Correlation

Tables 3 and 4 display the list of applicant and interviewer cues significantly correlated with the hirability impression obtained from annotations of the full interview. Note that only cues showing significant correlations ($p < .05$) are displayed. Please refer to the main paper for the discussion about these results.

Table 3: Applicant nonverbal cues extracted from thin slices (whole-slices) and from the full interview significantly correlated with the hirability impressions annotated from the full interview ($p < .05$, $\dagger p < .005$). $N = 62$.

Feature	Full	Slice 1	Slice 2	Slice 3	Slice 4	Slice 5	Slice 6	Slice 7	Slice 8
<i>Applicant prosodic features (audio):</i>									
Applicant pitch std	-0.44 [†]	-0.39 [†]	-0.29	-0.42 [†]	-0.48 [†]	-0.44 [†]	-0.45 [†]	-0.39 [†]	-0.41 [†]
Applicant pitch median	0.28	0.25	0.28	0.28	0.28	0.26	0.29	0.28	0.31
Applicant pitch lower quartile	0.39 [†]	0.35	0.36 [†]	0.41 [†]	0.40 [†]	0.39 [†]	0.37 [†]	0.37 [†]	0.37 [†]
Applicant energy median	0.33	0.31	0.23	0.29	0.34	0.34	0.32	0.28	
Applicant energy lower quartile	0.37 [†]	0.43 [†]	0.36 [†]	0.35	0.35	0.34	0.33	0.35	
Applicant energy upper quartile	0.54 [†]	0.38 [†]	0.42 [†]	0.27	0.36 [†]	0.44 [†]	0.47 [†]	0.53 [†]	
Applicant voiced rate mean	0.50 [†]	0.30	0.43 [†]	0.32	0.30	0.42 [†]	0.44 [†]	0.44 [†]	0.29
Applicant voiced rate std	0.32	0.43 [†]	0.40 [†]	0.28	0.28	0.35	0.34	0.40 [†]	
Applicant voiced rate lower quartile	0.47 [†]	0.29	0.40 [†]	0.28	0.37 [†]	0.29	0.41 [†]	0.42 [†]	0.29
Applicant voiced rate upper quartile	0.48 [†]	0.45 [†]	0.39 [†]	-0.33	-0.37 [†]	-0.38 [†]	-0.37 [†]	-0.38 [†]	-0.38 [†]
Applicant voiced rate maximum	0.53 [†]	0.39 [†]	0.50 [†]	0.31	-0.31	-0.30	-0.30	-0.32	-0.30
<i>Applicant turn based features (audio):</i>									
Applicant # of turns	-0.58 [†]	-0.39 [†]	-0.38 [†]	-0.37 [†]	-0.38 [†]	-0.37 [†]	-0.33	-0.38 [†]	-0.38 [†]
Applicant speaking time	0.48 [†]	0.45 [†]	0.39 [†]	-0.26	-0.44 [†]	-0.35	-0.33	-0.33	-0.27
Applicant average turn duration	0.53 [†]	-0.31	-0.31	-0.39 [†]	-0.46 [†]	-0.37 [†]	-0.46 [†]	-0.40 [†]	-0.26
Applicant maximum turn duration	-0.50 [†]	-0.39 [†]	-0.43 [†]	-0.41 [†]	-0.37 [†]	-0.48 [†]	-0.38 [†]	-0.43 [†]	-0.44 [†]
Applicant number of pauses	-0.58 [†]	-0.41 [†]	-0.41 [†]	-0.41 [†]	-0.37 [†]	-0.31	-0.29	-0.31	-0.46 [†]
Number of silent segments	-0.58 [†]	-0.32	-0.33	-0.30	-0.32	-0.32	-0.32	-0.35	-0.36 [†]
Total silent time	-0.58 [†]	-0.35	-0.43 [†]	-0.35	-0.38 [†]	-0.43 [†]	-0.37 [†]	-0.37 [†]	-0.45 [†]
Number of overlapping segments	-0.45 [†]	-0.35	-0.46 [†]	-0.35	-0.39 [†]	-0.43 [†]	-0.37 [†]	-0.37 [†]	-0.44 [†]
Total overlapping time	-0.45 [†]	-0.32	-0.32	-0.32	-0.32	-0.32	-0.32	-0.38 [†]	-0.44 [†]
Applicant number of short utterances	-0.45 [†]	-0.35	-0.45 [†]	-0.35	-0.38 [†]	-0.43 [†]	-0.37 [†]	-0.37 [†]	-0.45 [†]
Applicant total short utterance time	-0.45 [†]	-0.32	-0.32	-0.32	-0.39 [†]	-0.43 [†]	-0.37 [†]	-0.38 [†]	-0.44 [†]
Applicant # of audio back-channels	-0.32	-0.32	-0.28	-0.28	-0.27	-0.27	-0.27	-0.32	-0.32
Applicant audio back-channel time	-0.32	-0.29	-0.26	-0.26	-0.26	-0.26	-0.26	-0.36 [†]	-0.36 [†]
Applicant WMEI features (visual):									
Applicant WMEI vertical center of mass	-0.29	-0.28	-0.28	-0.32	-0.26	-0.26	-0.26	-0.34	-0.34
Applicant WMEI non-zero ratio	-0.28	-0.29	-0.29	-0.28	-0.28	-0.28	-0.28	-0.30	-0.30
Applicant nod-based features (visual):									
Applicant number of head nods	0.26	-0.26	-0.26	-0.29	-0.29	-0.29	-0.29	-0.33	-0.33
Applicant nodding time	0.26	-0.29	-0.29	-0.30	-0.28	-0.28	-0.28	-0.27	-0.27
Applicant number of visual back-channel events	0.26	-0.26	-0.26	-0.29	-0.29	-0.29	-0.29	-0.31	-0.31
Applicant visual back-channeling time	0.26	-0.29	-0.29	-0.30	-0.30	-0.30	-0.30	-0.27	-0.27
Applicant number of noddings while speaking events	0.26	-0.26	-0.26	-0.29	-0.29	-0.29	-0.29	-0.31	-0.31
Applicant head motion features (visual):									
Applicant median horizontal head motion	0.29	-0.31	-0.31	-0.26	-0.26	-0.26	-0.26	-0.34	-0.34
Applicant maximum horizontal head motion	0.31	0.31	0.26	0.33	0.29	0.31	0.31	0.37 [†]	0.37 [†]
Applicant mean vertical motion	0.40 [†]	0.40 [†]	0.40 [†]	0.37 [†]	0.37 [†]	0.35	0.35	0.44 [†]	0.44 [†]
Applicant median vertical head motion	-0.25	-0.25	-0.25	-0.30	-0.30	-0.30	-0.30	-0.34	-0.34
Applicant maximum vertical head motion	0.31	0.31	0.38 [†]	0.29	0.26	0.25	0.25	0.31	0.31
Applicant mean head motion magnitude	0.31	0.31	0.38 [†]	0.26	0.26	0.25	0.25	0.30	0.31
Applicant maximum head motion magnitude	0.31	0.31	0.38 [†]	0.26	0.26	0.25	0.25	0.31	0.31

Table 4: Interviewer nonverbal cues extracted from thin slices (whole-slices) and from the full interview significantly correlated with the hirability impressions annotated from the full interview ($p < .05$, $\dagger p < .005$). $N = 62$.

Feature	Full	Slice 1	Slice 2	Slice 3	Slice 4	Slice 5	Slice 6	Slice 7	Slice 8
<i>Interviewer prosodic features (audio):</i>									
Interviewer pitch std	-0.42 [†]	-0.33	-0.36 [†]	-0.30	-0.26	-0.32	-0.32	-0.28	0.32
Interviewer pitch lower quartile		-0.30	0.34						
Interviewer pitch upper quartile			-0.26						
Interviewer pitch maximum			0.25						
Interviewer energy median	0.30	0.27	0.37 [†]						
Interviewer energy lower quartile									
Interviewer voiced rate mean									
Interviewer voiced rate median									
Interviewer voiced rate lower quartile								0.27	
<i>Interviewer turn-based features (audio):</i>									
Interviewer number of turns	-0.34	-0.30	-0.34	-0.33	-0.30	-0.30	-0.30	-0.34	-0.30
Interviewer speaking time	-0.33	-0.31	-0.31	-0.33	-0.33	-0.32	-0.32	-0.33	-0.34
Interviewer average turn duration	-0.32	-0.31	-0.28	-0.32	-0.32	-0.31	-0.32	-0.32	-0.31
Interviewer maximum turn duration	-0.37 [†]	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29	-0.33
Interviewer number of pauses	-0.50 [†]	-0.43 [†]	-0.39 [†]	-0.46 [†]	-0.37 [†]	-0.46 [†]	-0.46 [†]	-0.40 [†]	-0.44 [†]
Number of silent segments	-0.58 [†]	-0.41 [†]	-0.41 [†]	-0.37 [†]	-0.48 [†]	-0.38 [†]	-0.42 [†]	-0.43 [†]	-0.46 [†]
Total silent time	-0.32	-0.32	-0.31	-0.29	-0.32	-0.32	-0.32	-0.31	-0.36 [†]
Number of overlapping segments	-0.33	-0.33	-0.30	-0.30	-0.32	-0.32	-0.32	-0.35	-0.37 [†]
Total overlapping time	-0.33	-0.33	-0.33	-0.36 [†]	-0.36 [†]	-0.39 [†]	-0.39 [†]	-0.31	-0.42 [†]
Interviewer number of short utterances	-0.37 [†]	-0.33	-0.34	-0.37 [†]	-0.37 [†]	-0.39 [†]	-0.39 [†]	-0.32	-0.41 [†]
Interviewer short utterance time	-0.33	-0.33	-0.30	-0.30	-0.32	-0.32	-0.32	-0.28	
Interviewer number of audio back-channel events	-0.33	-0.33	-0.30	-0.30	-0.33	-0.33	-0.33	-0.30	
Interviewer audio back-channeling time									
<i>Interviewer WMEI features (visual):</i>									
Interviewer WMEI upper quartile								-0.26	
Interviewer nod-based features (visual):									
Interviewer number of head nods	0.35	-0.32	-0.29	-0.29	-0.29	-0.29	-0.29	-0.27	-0.31
Interviewer nodding time	0.42 [†]	-0.32	-0.29	-0.29	-0.29	-0.29	-0.29	-0.29	-0.30
Interviewer number of visual back-channel events	0.51 [†]	-0.32	-0.27	-0.27	-0.26	-0.26	-0.26	-0.26	-0.28
Interviewer visual back-channeling time	0.54 [†]	-0.31	-0.28	-0.28	-0.26	-0.26	-0.26	-0.27	-0.27
Interviewer number of nodding while speaking events	-0.35	-0.33	-0.32	-0.32	-0.32	-0.32	-0.32	-0.32	
Interviewer nodding while speaking time									
<i>Interviewer head motion features (visual):</i>									
Interviewer mean horizontal head motion	0.26	0.27	0.27	0.27	0.26	0.26	0.26	0.26	0.31
Interviewer median horizontal head motion	0.31	0.35	0.37 [†]	0.37 [†]	0.33	0.33	0.33	0.31	
Interviewer mean vertical head motion									
Interviewer median vertical head motion									
Interviewer mean head motion magnitude	0.28	0.26	0.33	0.33	0.28	0.28	0.28	0.28	