Article

How Anxiety State Influences Speech Parameters: A Network Analysis Study from a Real Stressed Scenario

Qingyi Wang [1,2], Feifei Xu [1], Xianyang Wang [2], Shengjun Wu [2], Lei Ren [3,4,*] and Xufeng Liu [2,*]

[1] School of Basic Medicine, Air Force Medical University, Xi'an 710032, China; jussy@126.com (Q.W.); 15102931199@163.com (F.X.)

[2] School of Military Medical Psychology, Air Force Medical University, Xi'an 710032, China; wangxianyang_1999@163.com (X.W.); wushj@fmmu.edu.cn (S.W.)

[3] Military Psychology Section, Logistics University of PAP, Tianjin 300309, China

[4] Military Mental Health Services & Research Center, Tianjin 300309, China

[*] Correspondence: rl_fmmu@163.com (L.R.); lllfmmu@fmmu.edu.cn (X.L.)

Abstract: Background/Objectives: Voice analysis has shown promise in anxiety assessment, yet traditional approaches examining isolated acoustic features yield inconsistent results. This study aimed to explore the relationship between anxiety states and vocal parameters from a network perspective in ecologically valid settings. **Methods**: A cross-sectional study was conducted with 316 undergraduate students (191 males, 125 females; mean age 20.3 ± 0.85 years) who completed a standardized picture description task while their speech was recorded. Participants were categorized into low-anxiety (n = 119) and high-anxiety (n=197) groups based on self-reported anxiety ratings. Five acoustic parameters—jitter, fundamental frequency (F0), formant frequencies (F1/F2), intensity, and speech rate—were analyzed using network analysis. **Results**: Network analysis revealed a robust negative relationship between jitter and state anxiety, with jitter as the sole speech parameter consistently linked to state anxiety in the total group. Additionally, higher anxiety levels were associated with a coupling between intensity and F1/F2, whereas the low-anxiety network displayed a sparser organization without intensity and F1/F2 connection. **Conclusions**: Anxiety could be recognized by speech parameter networks in ecological settings. The distinct pattern with the negative jitter-anxiety relationship in the total network and the connection between intensity and F1/2 in high-anxiety states suggest potential speech markers for anxiety assessment. These findings suggest that state anxiety may directly influence jitter and fundamentally restructure the relationships among speech features, highlighting the importance of examining jitter and speech parameter interactions rather than isolated values in speech detection of anxiety.

Keywords: anxiety assessment; speech parameters; network analysis; ecological validity; psychological stress

1. Introduction

The rapid pace and high-pressure demands in contemporary society can easily induce emotional responses such as the anxiety state in individuals [1,2]. These negative emotions may cause impaired decision-making [3], reduced work efficiency [4], and strained interpersonal relationships [1], ultimately impacting overall productivity and an individual's well-being [5]. These emotions are largely driven by the persistent stressors of daily life [2]. Distinct from chronic anxiety disorders, the anxiety state is a transient

emotional response that individuals experience when exposed to situational stressors, such as stringent deadlines, performance evaluations, and social pressures [6,7]. Despite this growing prevalence, the methods for its assessment remain subjective and time-consuming [8,9]. Traditional evaluation of anxiety has predominantly relied on self-report questionnaires, clinical interviews, and behavioral observations, which, while providing valuable insights into subjective experiences, are susceptible to reporting biases and recall limitations [10,11]. These conventional approaches are typically supplemented by physiological measurements, including cardiovascular monitoring, blood pressure measures, and cortisol sampling, which offer objective biological markers of stress responses [12,13]. Contemporary wearable devices enable continuous tracking of various physiological parameters, including heart rate variability or electrodermal activity [14]; however, these measurements require direct contact between the sensing elements and the skin surface.

Speech has in recent years arisen as a particularly promising approach for anxiety detection, with its non-invasive quality, remote measurement capability, and potential for continuous monitoring [15,16]. It could reflect psychological states through synchronizing muscles across respiratory, phonatory, and supralaryngeal organs, working in line with cognitive and emotional processing [17,18]. The intricate interplay between these systems makes speech particularly valuable as a biomarker, as it can capture significant motor, cognitive, and behavioral changes associated with mental health conditions [19,20].

Individuals with anxiety disorders exhibit heightened activation of the sympathetic nervous system under stress [21,22], which directly influences speech production through various pathways [23]. This manifests in different features, including speech rates, increased disturbances, and hesitations [8,24,25], which could be measured by speech parameters such as fundamental frequency, formant frequencies, and various spectral properties [9,26]. Most research studies on anxiety have reported increased F0 in anxious individuals [8,24]. Besides F0, parameters such as jitter, shimmer, and pause patterns [9,14] and speech rate [27] reveal further complexity in anxiety-induced vocal expression. Anxiety may further impair articulatory precision, reducing vowel clarity and altering formant bandwidths as a result of stress-induced physiological arousal [28–30]. The anxiety state, as a transient response to situational stressors, often correlates with vocal parameters, including elevated fundamental frequency (F0), increased intensity, accelerated speech rate, and decreased duration [9,14,31–33].

Studies have been performed on speech parameters in anxiety, highlighting key sensitive indicators including the following: (1) Fundamental frequency (F0), which represents the rate of vocal cord vibration and corresponds to perceived pitch, has been consistently identified as a primary indicator of stress. Multiple studies have demonstrated a reliable increase in F0 under anxious or stressed conditions [15,34]. This elevation in F0 has been attributed to increased muscle tension and respiratory changes during stress responses [35]. (2) Formant frequencies, particularly the first (F1) and second (F2) formants, serve as crucial indicators of vocal tract resonance and articulation patterns. Research has shown that stress-induced changes in muscle tension and breathing patterns can affect formant characteristics [30]. The F1/F2 ratio has emerged as a potentially valuable metric, though individual variations in these parameters suggest speaker-specific patterns rather than universal trends [14,36]. (3) Jitter, which quantifies cycle-to-cycle frequency variation in vocal fold vibration, has demonstrated mixed results in anxiety detection. While some studies report decreased jitter under stress conditions [37], others have found more variable patterns, suggesting that jitter's relationship with anxiety may be modulated by individual differences and specific stressor types [15]. (4) Intensity, measured as the amplitude or loudness of speech signals, typically shows elevation under an anxious state, reflecting increased subglottal pressure and muscular tension [38]. However, the magnitude of these changes can vary significantly based on the nature and intensity of the stressor. (5) Speech

rate, often quantified through mean voiced segment length, indicates temporal features of stress-affected speech production. Stress typically leads to alterations in speaking patterns, including pause and segment duration [35]. Additionally, spectral features (MFCC, LPCC) and the harmonics-to-noise ratio (HNR) also tap into the acoustic complexity associated with stress, yet HNR sensitivity appears more consistent in physical stress (e.g., workouts) than in tasks involving heavy cognitive load or psychological stress [37–40]. Overall, these findings support the notion that anxiety-driven physiological arousal alters speech production through multiple, interrelated pathways, positioning acoustic features as viable indicators of psychological stress in various contexts.

To date, much of the existing research on anxious speech has been conducted in controlled laboratory settings, with limited ecological validity in the findings [28]. Real-world scenarios, such as those involving social evaluation or time pressure, presumably would be more ecologically valid for investigating anxiety-related speech changes. Moreover, previous studies have primarily focused on isolated speech parameters, overlooking the complex interrelations between these features [25,41]. For example, while F0 and jitter are often studied independently, their interactions with formants, intensity, and speech rate under anxiety state are poorly understood.

Network analysis, which is one of the state-of-the-art tools to identify and analyze the pattern of statistical association in multivariate data, has achieved exponential development in the field of psychological science [42]. Networks include nodes and edges. Nodes represent the variables of study and edges represent connections between nodes. Compared with traditional statistical models, network analysis offers the following methodological advantages for investigating the relationships among anxiety state and related speech parameters: (1) Visualization of systemic dynamics. It illustrates direct edges, indirect pathways, and central nodes in a unified framework, mitigating the reductionism and isolated focus of traditional methods [43]. (2) Group-level network comparison. Rather than treating parameters as independent units, such as t-tests on speech rate, network comparison tests reveal differences in edge connections, node centrality, and overall connectivity between low- and high-anxiety states [44]. (3) Statistically robust edge estimation. Within a network, edges are usually statistically depicted using regularized partial correlations. These correlations, obtained after controlling for other variables and employing statistical regularization techniques, represent purer, more restrained, and interpretable connections among multivariate data [45]. By integrating these advantages, network analysis offers a holistic, data-driven framework that not only pinpoints "hub" speech parameters but also uncovers subtle systemic dysregulations that may be overlooked by traditional methods.

Therefore, this study aims to examine the complex impacts of the anxiety state on speech parameters using network analysis in a real stressed scenario. Specifically, the study has two primary objectives: (1) to investigate the complex associations between the anxiety state and speech parameters, with a focus on both direct and indirect relationships within the network; and (2) to explore how different levels of an anxiety state influence the patterns of associations between speech parameters.

Based on existing evidence showing that the anxiety state is associated with changes in vocal properties, we hypothesize that (1) speech parameters will exhibit a direct association with the anxiety state, and (2) higher levels of anxiety will fundamentally alter the network relationships among key speech parameters compared to lower anxiety states.

2. Materials and Methods

2.1. Settings and Participants

Participants were recruited from second-year undergraduate students at a large public university in Xi'an City, China. Participants were selected using a restricted homogeneous sampling method. Out of 569 students in the same grade, over 334 subjects were chosen based on their similar CET-4 test scores (a standardized English proficiency test in China; M = 499, SD = 32.85) to ensure comparable language proficiency across participants. A total of 316 valid participants were ultimately included in the analysis (191 males [60.4%] and 125 females [39.6%]). The age of participants ranged from 19 to 22 years old, with a mean age of 20.3 (±0.85) years. All participants provided written informed consent prior to participation.

The oral examination was conducted by course instructors as part of the curriculum. Following the image description task, participants were immediately instructed to self-rate their anxiety levels experienced during the task. The research assistants provided clear instructions, asking participants to "reflect on your feelings during the image description task and rate your anxiety level based on your immediate experience". The self-rating was conducted using a 4-point scale (0 = no anxiety, 1 = mild anxiety, 2 = moderate anxiety, 3 = severe anxiety). This immediate assessment timing was crucial to capture the authentic anxiety-state levels while the experience was still fresh in participants' memory, minimizing potential recall bias. Based on these ratings, participants were categorized into two groups: the low-anxiety group (ratings of 0–1; n = 119) and the high-anxiety group (ratings of 2–3; n = 197). This self-assessment approach aligns with prior research emphasizing the subjective nature of the anxiety state [46,47]. Following the assessment, the university counseling service were available for students who reported high anxiety levels. This study was conducted in accordance with the Declaration of Helsinki and received approval from the Ethics Committee of Xijing Hospital (KY20242053-C-1).

2.2. Materials and Procedure

This study utilized a cross-sectional design to explore differences in speech parameters between high- and low-anxiety individuals under a naturalistic stress-inducing condition. The anxiety-eliciting task was a foreign language oral examination, where participants were required to describe a thematic cartoon in English. This foreign language exam is an established ecological stressor widely recognized for its ability to provoke anxiety in second-language learners [48,49] because of the performance-related stress, fear of failing linguistic assessments, and anticipation of negative evaluations from proficient instructors [49–51]. The assessment consisted of a standardized picture description task featuring a thematic cartoon (topic: online and offline life) selected from a validated resource for formal language assessment materials. The instructions and interface of the online language testing platform are presented in Supplementary Figure S4. Previous psychophysiological studies have documented significant stress responses in such contexts, particularly among learners in high-stakes educational environments, with consistent elevations in both subjective anxiety measures and objective stress indicators [52,53].

Recordings were conducted in the university language labs with the ambient noise level maintained below 40 dB. The lab has a typical office acoustical setup with moderate reverberation time (approximately 0.4–0.5 s). Participants completed the oral exam task in a computer laboratory equipped with standardized workstations. Each participant was provided with a SANAKO SLH-07 headset featuring an integrated unidirectional microphone (frequency response: 40 Hz–16 kHz) to ensure consistent audio recording quality across all participants. The oral examination was conducted through an online English-language testing platform (https://www.tsinghuaelt.com/ (accessed on 18 January 2024))

designed for Chinese college curricula, with identical audio settings across all workstations. The microphone functionality was tested on the platform's recording test page before beginning the oral exam (see Supplementary Figure S5).

The task instructions were presented uniformly through the testing platform interface. Students were required to talk about the picture displayed on the screen and were given in 45 s for preparation and 3 min to talk about it. After completing the task, participants were prompted to rate their anxiety state using the 4-point scale described earlier.

2.3. Speech Feature Extraction

Speech recordings were processed to extract five key acoustic parameters: jitter, fundamental frequency (F0), formants (F1/F2), intensity, and speech rate. These parameters were selected based on their established relevance to anxiety-related changes in speech [15,34,35].

The acoustic parameters analyzed in this study were defined as follows: Jitter refers to the cycle-to-cycle variation in fundamental frequency, which may indicate vocal instability and stress [23]. F0 reflects emotional arousal and tension [26]. F1/F2, the ratio of the first and second formants, represents interactions between pitch and vocal tract resonance and serves as a key indicator of articulatory precision and speech clarity [14]. Intensity captures the loudness of speech, measured in decibels (dB), and is often associated with vocal effort [54]. Speech rate measures the number of spoken syllables per second and reflects overall speech fluency and pacing [55].

Audio recordings were converted to a standardized format (16-bit WAV, 44.1 kHz) and trimmed to remove silence by Adobe Audition. Acoustic parameters were extracted using the open-source Python library Librosa (v0.10.0) and custom routines. Signals were pre-processed with a Hamming window (512 samples ≈ 23 ms at 22.05 kHz), 50% overlap (hop = 256 samples), and 0.97 pre-emphasis. F0 mean was extracted from indices 18/17 of the hsf array (Praat-aligned via Librosa). Formants F1/F2 were derived via 12th-order LPC analysis averaged across voiced segments using np.nanmean. Jitter followed the Praat equation. Intensity reflected RMS amplitude averaged over frames. Speech rate used Whisper ("base") transcriptions.

2.4. Data Analysis

We utilized SPSS software (version 27.0) to perform the descriptive statistical analysis. An independent samples t-test was conducted to compare differences between the high- and low-anxiety groups regarding various acoustic parameters, including intensity, speech rate, F1/F2, F0, jitter, and state anxiety. Before constructing three networks (i.e., total, low-anxiety, and high-anxiety networks), the goldbricker function in the R package network tools was applied to determine the redundant nodes [56], with no potential node redundancy screened out in the three networks.

The three networks in this study were created using graphical Least Absolute Shrinkage and Selection Operator (gLASSO) combined with the Extended Bayesian Information Criterion (EBIC) [45]. Within each network, the edges reflect the partial (Spearman) correlation between two paired nodes, accounting for the influence of other nodes [45,57]. The gLASSO technique punishes small correlation coefficients to zero and results in a sparser and more interpretable network [45,58]. The EBIC tuning parameter (gamma) was adjusted to 0.5. The gamma controls the severity of the model selection, with higher values indicating that simpler models are preferred [45]. When gamma is set to 0.5, most spurious edges can be avoided [59]. Therefore, this more conservative approach makes the results of network structure more likely to be stable and reproducible. The EBICglasso function introduces gamma = 0.5 as "generally a good choice" and sets it as the default value

[45,59]. The network visualization was generated using the Fruchterman–Reingold algorithm, through which nodes with weak and sparse connections are located on the periphery of the network, while nodes with strong and numerous connections tend to appear near the center of the network [60,61].

Expected influence, the sum value of all edges connecting to a specific node, was calculated for each node within low-anxiety and high-anxiety networks via R-package qgraph [61]. This indicator, when compared with the traditional centrality index (e.g., strength centrality), is more appropriate for networks that include both positive and negative edges [62]. Higher expected influence value of a node indicates its greater importance in the network [62,63].

To identify the accuracy of edge weights, a 95% confidence interval (1,000 bootstrap samples) was plotted for each edge [64]. The stability of expected influence was determined by testing the correlation stability (CS) coefficient using the case-dropping bootstrap method (1000 bootstrap samples). The CS coefficient measures the maximum proportion of data that can be dropped to retain, with 95% certainty, a correlation of at least 0.7 with the centralities of the original network [64]. As recommended by Epskamp et al., it is considered ideal for the CS coefficient to be higher than 0.5 and not lower than 0.25 [64]. Additionally, we carried out bootstrapped difference tests (1000 bootstrap samples) for edge weights and expected influence. These analyses were performed using the R-package bootnet [64].

To investigate differences between low-anxiety and high-anxiety network characteristics, the R-package NetworkComparisonTest was employed with 1000 permutations [44]. We primarily concentrated on three key network characteristics: global expected influence, edge weights, and node expected influences. Given the exploratory nature of this study, no corrections were utilized for multiple comparisons [44,65].

3. Results

Table 1 presents the descriptive statistics for the anxiety state and acoustic parameters across the total sample, as well as the low- and high-anxiety groups. Self-reported anxiety differed significantly between groups, with the low-anxiety group reporting markedly lower levels compared to the high-anxiety group ($p < 0.001$, Cohen's d = −3.731). This large effect size highlights a clinically meaningful distinction in perceived anxiety levels between the groups.

For acoustic parameters, jitter was slightly higher in the low-anxiety group compared to the high-anxiety group ($p < 0.05$, Cohen's d = 0.229), and F0 was slightly lower in the low-anxiety group than in the high-anxiety group ($p < 0.05$, Cohen's d = −0.232). Both differences were statistically significant but exhibited small effect sizes.

No statistically significant differences were observed for formant ratio (F1/F2), intensity, or speech rate ($p > 0.2$), and these variables demonstrated negligible effect sizes.

Table 1. Anxiety state and acoustic variables across total, low-anxiety, and high-anxiety groups.

Variables	Total Group (Mean ± SD)	Low-Anxiety Group (Mean ± SD)	High-Anxiety Group (Mean ± SD)	Low-Anxiety Group vs. High-Anxiety Group (p-Value)	Cohen's d
Anxiety state	1.83 ± 1.01	0.70 ± 0.46	2.51 ± 0.50	0.000 **	−3.731
Jitter	0.50 ± 0.07	0.51 ± 0.06	0.49 ± 0.08	0.040 *	0.229
F0	181.88 ± 18.53	179.20 ± 18.70	183.49 ± 18.29	0.046 *	−0.232
F1/F2	0.25 ± 0.05	0.25 ± 0.06	0.25 ± 0.05	0.212	0.145
Intensity	67.54 ± 5.64	67.78 ± 5.41	67.40 ± 5.78	0.561	0.068
Speech rate	1.26 ± 0.57	1.26 ± 0.61	1.26 ± 0.55	0.999	0.000

Variables of Significance:*$p \le 0.05$, **$p \le 0.001$.

Figure 1 illustrates the network structure of the anxiety state and various speech-related indicators in the total group. It is evident that the anxiety state is negatively connected to only one speech indicator—jitter (edge weight = −0.17). The connections between the speech indicators themselves are generally strong and predominantly positive. Among the speech indicators, three positive connections are, respectively, speech rate–intensity (edge weight = 0.28), intensity–F1/2 (edge weight = 0.26), and F1/2–F0 (edge weight = 0.24). There exists only one strong negative connection between jitter and F0 (edge weight = −0.39). The bootstrapped 95% confidence interval indicates that the edge weights are relatively accurate (Figure S1a in the Supplementary Material). Figure S1b shows the difference test results of all edges.
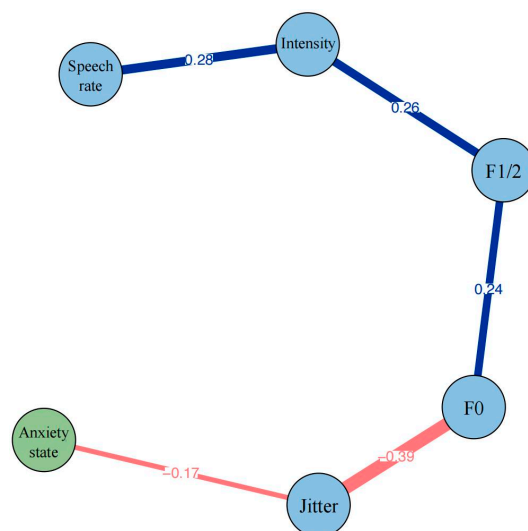

Figure 1. Network structure of anxiety state and speech indicators in total group. Blue edges represent positive correlations and red edges represent negative correlations. The thickness of the edge reflects the magnitude of the correlation. The value on the edge represents the edge weights of two corresponding nodes.

In the speech indicator network of the high-anxiety-state group (see upper left of Figure 2), the pattern of connections between speech indicators is very similar to that in Figure 1. Three positive connections are, respectively, speech rate–intensity (edge weight = 0.28), intensity–F1/2 (edge weight = 0.33), and F1/2–F0 (edge weight = 0.20). Only one strong negative connection is between jitter and F0 (edge weight = −0.39). In the speech indicator network in the low-anxiety-state group (see lower left of Figure 2), except for the absence of a connection between intensity and F1/2, the connection patterns between the other indicators remain similar to the high-anxiety group. Bootstrapped 95% confidence intervals indicate that the edge weights in the two samples are relatively accurate (Figure S2a,c in the Supplementary Material). Figure S2b,d show the edge weights' difference test results.
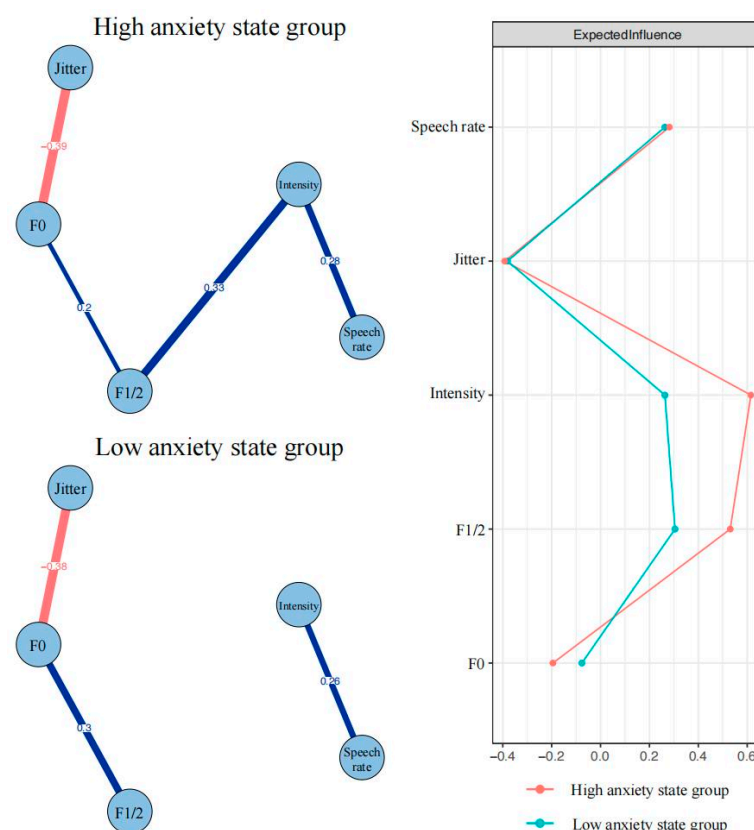
Figure 2. Speech indicator networks in high- and low-anxiety-state groups using average layout and bridge centrality plots. Blue edges represent positive correlations and red edges represent negative correlations. The thickness of the edge reflects the magnitude of the correlation. The value on the edge represents the edge weights of two corresponding nodes.

The expected influence results for the high-anxiety- and low-anxiety-state groups are plotted in the right section of Figure 2. In the high-anxiety-state network, intensity has the strongest expected influence (raw value = 0.62). In the low-anxiety-state network, F1/2 has the strongest expected influence (raw value = 0.30). The CS coefficients of expected influences for high-anxiety-state and low-anxiety-state networks are 0.67 and 0.36, respectively (Figure S3a,c in the Supplementary Material). Figure S3b,d show the nodes expected influences' difference test results.

The results of the network comparison test reveal that only one connection among speech indicators exhibit significant differences: F1/2-intensity (p = 0.02). The node expected influences of two groups show no significant differences. In addition, the global expected influence of two groups also show no significant differences (high-anxiety-state group = 0.42; low-anxiety-state group = 0.19; S = 0.23, p = 0.43).

4. Discussion

This study employed network analysis to explore the impacts of anxiety state on speech parameters in ecologically valid settings. Two distinct networks were constructed. The first network elucidated the relationships between anxiety state and speech parameters, with particular focus on the direct associations and transmission effects between parameters. The second network identified how different levels of anxiety state influence the association patterns among speech parameters.

From a global view, the network structure demonstrated a clear, chain-like arrangement of speech parameters, with jitter emerging as the only parameter directly connected to the anxiety state. This finding may suggest the role of jitter as the most critical acoustic

indicator of anxiety responses under stress. Specifically, the correlation between the anxiety state and jitter indicates that increased anxiety levels may disturb the vocal stability and trigger alterations in pitch. This finding aligns with previous research suggesting that anxiety-induced physiological changes affect laryngeal muscle tension and respiratory control [66]. Notably, the effect of anxiety on jitter is not universally observed. Some studies have documented increased jitter in response to stress [9,14,67], often attributed to heightened laryngeal and cricothyroid tension leading to less stable vocal fold vibration. Other research, however, includes findings similar to ours, where jitter decreases under real-world stress conditions such as exam or emergencies [37,68]. The reduced jitter in these scenarios could be justified by the cognitive load and emotional load induced alone or in combination [35]. And it could also be attributed to the compensatory vocal mechanisms under anxiety conditions, suggesting heightened psychological arousal may lead to increased laryngeal muscle tension, consequently leading to more regular vocal fold vibration. Alternatively, given the dynamic nature of anxiety during a speech task, speakers may experience peak anxiety at the beginning of a task, which may gradually diminish as they adapt to the stressor over time. This temporal variation in anxiety could influence jitter differently at different phases of the task. As such, if jitter reduction is driven by compensatory mechanisms, this effect may emerge later in the task as speakers adapt to the stressors.

From the perspective of individual differences, these contradictory patterns may reflect variability in speech motor control or baseline vocal tension. When in a stressed state, jitter could be caused by either small variations or asymmetries in the cricothyroid muscle tension [69] and/or fluctuations in subglottal pressure [70] and/or perturbations in the mucous of the vocal folds [71]. Speakers with greater baseline vocal tension may show a decrease in jitter as a form of physiologically "locked" vocal fold vibration, whereas those with lower baseline tension might exhibit an increase in jitter when confronted with stress. So, baseline vocal tension could be measured in the future by surface electromyography of the vocal tract, or general speech motor control could be pre-assessed to elucidate the individual differences in jitter responses under anxiety.

Most importantly, our research extends beyond the confines of the lab, offering greater ecological validity by examining stress responses in more naturalistic settings. Previous network analysis study conducted in a controlled laboratory setting identified jitter as the only speech parameter directly connected to self-reported negative affect [14]. Consistently, we also found jitter to be the only speech parameter directly linked to stress responses, reinforcing its potential as a robust acoustic biomarker of stress. These findings suggest that jitter holds significant promise as a reliable indicator of stress whether in controlled experimental environments or in daily life.

Furthermore, jitter also correlates inversely with F0, possibly acting as a "bridge" variable in transmitting anxiety effects onto other vocal parameters. A direct connection with F0 was expected, as jitter by definition refers to the variations that occur in F0. Previous research mostly showed that increased fundamental frequency (F0) is accompanied by elevated cycle-to-cycle frequency perturbation (jitter) under stress conditions [37,72,73]. As Giddens et al. reported, increased vocal fold tension under stress typically raises F0 while also leading to more irregular vocal fold vibration patterns, resulting in increased jitter [15]. Interestingly, our finding implies that despite the stress-induced elevation in pitch, speakers maintain or even enhance their vocal stability, possibly as a physiological adaptation to maintain speech clarity under pressure. Together, this sequential organization suggests a hierarchical relationship in how anxiety manifests through various vocal parameters and potentially reflects the intricate mechanisms underlying speech production under stress.

The total group network analysis also showed a positive interplay among core speech parameters, where F0, formant frequencies (F1/2), intensity, and speech rate exhibited

positive interconnections, suggesting a coherent vocal pattern that might collectively shift when speaking under challenging conditions. From a physiological standpoint, as F0 rises, speakers often inadvertently increase the intensity and modify the articulation captured by F1/2, which can also interact with the rate of delivery. Specifically, the positive correlation between F0 and F1/2 suggests that as speakers raise their pitch under stress, they also exhibit higher formant frequencies. This relationship likely reflects complex changes in vocal tract configuration under stress, as documented by Scherer [74] and Murray et al. [75]. The positive relationship between F1/2 and intensity indicates that speakers who produce higher formant frequencies also tend to speak louder under stress, a pattern consistent with acoustic theory and supported by studies of stress-induced vocal changes, noting that stress-induced physiological changes often result in simultaneous increases in vocal intensity and articulatory effort [26]. The moderate positive correlation between intensity and speech rate further highlights the possible activation effect, wherein speakers under stress tend to speak both louder and slightly faster. Stress-induced arousal typically results in increased respiratory drive and heightened vocal cord tension, which can raise vocal intensity while also promoting a faster articulation tempo, as supported by studies on emotional vocal expression [26,74]. This relationship is consistent with findings that stress can amplify vocal energy and motor coordination, driving dynamic changes across vocal parameters. Moreover, the observed interconnections among F0, F1/2, intensity, and speech rate align with Scherer's push effect theory [74], in which physiological arousal during emotional states initiates coordinated adaptations across multiple vocal subsystems. These speech changes arise from increased sympathetic nervous system activation, subsequently affecting breathing patterns, vocal fold tension, and articulatory precision, eventually leading to synchronized shifts in pitch, loudness, and speech pacing. Thus, this chain of relationships reflects how stress modifies vocal behavior holistically.

Most interestingly, the network analysis revealed distinct patterns of acoustic parameter connections between high- and low-anxiety groups, indicating that anxiety may fundamentally restructure the relationships among vocal parameters under stress. Similar to the total group network, the high-anxiety network exhibited greater association among acoustic parameters, with multiple strong connections forming a chain-like structure. This increased parameter coupling in the high-anxiety group aligns with Scherer's component process model [76] and Goberman et al.'s stress-induced vocal synchronization [23], suggesting that tighter coupling among acoustic parameters may serve as reliable biomarkers for anxiety states. The low-anxiety network displayed a sparse structure with fewer strong connections between parameters, which may suggest more relaxed vocal control compared to high-anxiety states [35]. This finding is consistent with previous research showing that neutral or low-arousal states typically demonstrate more independent variation in acoustic features [26]. This separation between intensity and F1/2 in the low-anxiety network indicates that speakers could maintain their natural articulation without intensity compensation, in contrast to the strong connection of F1/2 and intensity in the high-anxiety group, where speakers need to enhance their vocal intensity to compensate for the weak articulation. This pattern likely reflects what Lindblom termed "hypo-speech", where speakers optimize articulatory effort to balance communicative efficiency and minimal physiological cost within a given context [77]. Thus, these contrasting networks offer valuable diagnostic potential for anxiety detection. The presence or absence of intensity–articulation coupling may serve as a reliable marker for anxiety states, particularly in ecologically valid settings.

The expected influence analysis provides additional insights into the relative importance of different speech parameters within each anxiety-state network. In the high-anxiety state, intensity (EI = 0.62) emerged as the most central node, suggesting the dominant role of intensity in the high-anxiety group, which reflects the compensatory vocal

effort to maintain clarity despite stress-induced physiological arousal [74]. Conversely, in the low-anxiety state, F1/F2 (EI = 0.30) showed the strongest influence, indicating greater articulatory freedom when cognitive resources are not consumed by anxiety. This pattern aligns with Lindblom's hyper- and hypo-speech framework, suggesting that speakers optimize vocal effort relative to communicative demands [77]. Comparing the expected influence between two groups, the low-anxiety-state network displayed more evenly distributed expected influence values across F1/2, intensity, and speech rate. This balanced distribution may suggest that parameters operate with greater independence rather than being dominated by specific features. The difference in expected influence between anxiety states further supports our earlier observations about the distinctive network organizations under different anxiety conditions. However, the absence of significant differences in nodes' expected influence and global expected influence between groups suggests that while the organization of speech parameters differs between anxiety states, the overall influence structure remains relatively stable.

It is also worth noting that the use of English as a non-native language in our task likely amplified the anxiety triggered by the exam through foreign language anxiety (FLA) mechanisms [78], particularly affecting articulatory precision (F1/F2) and vocal effort (intensity) [79,80]. This aligns with FLA research showing that linguistic insecurity exacerbates vocal changes [81,82]. Non-native speakers often exhibit reduced articulatory control and increased vocal tension due to unfamiliar phonemes [83,84], which may explain the strong intensity–F1/F2 coupling observed in our high-anxiety network. However, the stability of jitter as a biomarker across anxiety states suggests some vocal stress responses transcend language-specific factors.

Previous research on speech biomarkers primarily focused on evaluating individual acoustic parameters in isolation. However, our study indicates the importance of the relationships between these parameters as additional potential biomarkers, providing a fresh perspective on stress-related vocal dynamics. This shift broadens the scope of traditional biomarker research by emphasizing the need to examine how vocal features interact within a speech parameter network rather than concentrating solely on individual values.

This study first employs network analysis of vocal parameters in an ecologically valid anxiety context, filling the gap in the existing anxiety research that has largely been restricted to laboratory environments. In addressing the first objective, we uncovered a network across anxiety state and speech parameters, which highlights the direct associations between jitter and anxiety state, and transmission effects among vocal features. Importantly, jitter is found to be the sole speech parameter consistently linked to the anxiety state across all networks, underscoring its potential as a robust acoustic biomarker for stress. For the second objective, we demonstrated that varying levels of the anxiety state change the association patterns among speech features fundamentally, manifesting as the absence of an intensity and F1/F2 connection. These findings could enhance voice-based anxiety screening tools by prioritizing jitter and intensity–F1/F2 coupling as biomarkers. For instance, mobile apps analyzing speech during telehealth consultations or workplace stress assessments could flag high-risk states in real time. Clinically, this could supplement traditional surveys with objective vocal metrics.

Several methodological limitations warrant consideration. First, by recruiting mainly non-native English speakers from a single university, our findings may not fully generalize to other linguistic or cultural settings. Second, we relied on a single speech task (picture description), which may not capture the full range of anxiety-inducing scenarios—such as public speaking or interpersonal conflict. Furthermore, our binary categorization of anxiety (low vs. high) could obscure more nuanced or non-linear trajectories, including the possibility that jitter initially increases at moderate anxiety levels but later stabilizes as speakers

adapt. Additionally, anxiety was measured solely via self-report, and the absence of physiological indicators (e.g., cortisol, heart rate variability) may limit the objectivity of anxiety assessment. Finally, to keep computational demands manageable, we focused on a limited set of vocal parameters, potentially missing other relevant acoustic features.

Therefore, future research should recruit participants from diverse populations to broaden the applicability of the findings. Studies should also examine additional stress-inducing contexts and analyze speech production at multiple time points to better capture the dynamic progression of anxiety and the potential emergence of compensatory mechanisms. Furthermore, regression modeling of anxiety could be incorporated to determine the trend of jitter response. Further, incorporating physiological indicators and additional acoustic parameters, such as the harmonics-to-noise ratio, could be used to examine detailed anxiety-related vocal changes. Finally, individual differences in speech motor control and baseline anxiety should be explored to uncover the mechanisms linking anxiety with vocal parameters.

5. Conclusions

This study provides insights into the relationship between anxiety states and speech parameters using network analysis in an ecologically valid setting. By examining the interconnections among vocal features, we identified jitter as the sole speech parameter consistently linked to the anxiety state across all networks, underscoring its potential as a robust acoustic biomarker for stress. Additionally, the coupling between intensity and F1/F2 in high-anxiety states highlights how anxiety fundamentally reorganizes vocal parameter relationships, reflecting compensatory mechanisms under stress. Our findings contribute to a more comprehensive understanding of how anxiety manifests through interconnected vocal parameters. We also demonstrate that the ecological validity of using real-world stressors, such as oral examinations, could capture authentic anxiety responses, which enhances the applicability of our results to real-life contexts. These findings have practical implications for developing automated, speech-based anxiety screening tools that prioritize jitter and intensity–F1/F2 coupling as key markers. This study may offer a fresh perspective for developing objective, non-invasive assessments of anxiety.

References

1. Association, A.P. Stress in America 2022: Concerned for the Future, Beset by Inflation. Available online: https://www.apa.org/news/press/releases/stress/2022/concerned-future-inflation (accessed on 12 December 2024).

2. Almeida, D.M.; Charles, S.T.; Mogle, J.; Drewelies, J.; Aldwin, C.M.; Spiro, A.; Gerstorf, D. Charting Adult Development through (Historically Changing) Daily Stress Processes. *Am. Psychol.* **2020**, *75*, 511–524. https://doi.org/10.1037/amp0000597.

3. Hartley, C.A.; Phelps, E.A. Anxiety and Decision-Making. *Biol. Psychiatry* **2012**, *72*, 113–118. https://doi.org/10.1016/j.biopsych.2011.12.027.

4. Ford, M.T.; Cerasoli, C.P.; Higgins, J.A.; Decesare, A.L. Relationships between Psychological, Physical, and Behavioural Health and Work Performance: A Review and Meta-Analysis. *Work Stress* **2011**, *25*, 185–204.

5. Chisholm, D.; Sweeny, K.; Sheehan, P.; Rasmussen, B.; Smit, F.; Cuijpers, P.; Saxena, S. Scaling-up Treatment of Depression and Anxiety: A Global Return on Investment Analysis. *Lancet Psychiatry* **2016**, *3*, 415–424.

6. Spielberger, C.D.; Sydeman, S.J. State-Trait Anxiety Inventory and State-Trait Anger Expression Inventory. In *The Use of Psychological Testing for Treatment Planning and Outcome Assessment*; Maruish, M.E., Ed.; Lawrence Erlbaum Associates, Inc.: Mahwah, NJ, USA, 1994; pp. 292–321.

7. Eysenck, M.W.; Derakshan, N.; Santos, R.; Calvo, M.G. Anxiety and Cognitive Performance: Attentional Control Theory. *Emotion* **2007**, *7*, 336.

8. Riad, R.; Denais, M.; de Gennes, M.; Lesage, A.; Cao, X.-N.; Mouchabac, S.; Bourla, A. Automated Speech Analysis for Risk Detection of Depression, Anxiety, Insomnia, and Fatigue: Algorithm Development and Validation Study. *J. Med. Internet Res.* **2024**, *26*, e58572.

9. Özseven, T.; Düğenci, M.; Doruk, A.; Kahraman, H.İ. Voice Traces of Anxiety: Acoustic Parameters Affected by Anxiety Disorder. *Arch. Acoust.* **2018**, *43*, 625–636. https://doi.org/10.24425/aoa.2018.125156.

10. Fink, G. *Stress: Concepts, Cognition, Emotion, and Behavior: Handbook of Stress*; Academic Press: Cambridge, MA, USA, 2016.

11. Monroe, S.M. Modern Approaches to Conceptualizing and Measuring Human Life Stress. *Annu. Rev. Clin. Psychol.* **2008**, *4*, 33–52. https://doi.org/10.1146/annurev.clinpsy.4.022007.141207.

12. Shields, G.S.; Slavich, G.M. Lifetime Stress Exposure and Health: A Review of Contemporary Assessment Methods and Biological Mechanisms. *Soc. Pers. Psychol. Compass* **2017**, *11*, e12335. https://doi.org/10.1111/spc3.12335.

13. Kirschbaum, C.; Hellhammer, D.H. Salivary Cortisol in Psychoneuroendocrine Research: Recent Developments and Applications. *Psychoneuroendocrinology* **1994**, *19*, 313–333. https://doi.org/10.1016/0306-4530(94)90013-2.

14. Kappen, M.; Hoorelbeke, K.; Madhu, N.; Demuynck, K.; Vanderhasselt, M.A. Speech as an Indicator for Psychosocial Stress: A Network Analytic Approach. *Behav. Res. Methods* **2022**, *54*, 910–921. https://doi.org/10.3758/s13428-021-01670-x.

15. Giddens, C.L.; Barron, K.W.; Byrd-Craven, J.; Clark, K.F.; Winter, A.S. Vocal Indices of Stress: A Review. *J. Voice* **2013**, *27*, 390.e21–390.e29. https://doi.org/10.1016/j.jvoice.2012.12.010.

16. Slavich, G.M.; Taylor, S.; Picard, R.W. Stress Measurement Using Speech: Recent Advancements, Validation Issues, and Ethical and Privacy Considerations. *Stress* **2019**, *22*, 408–413. https://doi.org/10.1080/10253890.2019.1584180.

17. Guenther, F.H. *Neural Control of Speech*; CogNet; The MIT Press: Cambridge, MA, USA, 2016; ISBN 978-0-262-33698-7.

18. Levelt, W.J.M. *Speaking: From Intention to Articulation*; The MIT Press: Cambridge, MA, USA, 1989.

19. Low, D.M.; Bentley, K.H.; Ghosh, S.S. Automated Assessment of Psychiatric Disorders Using Speech: A Systematic Review. *Laryngoscope Investig. Otolaryngol.* **2020**, *5*, 96–116. https://doi.org/10.1002/lio2.354.

20. Cummins, N.; Scherer, S.; Krajewski, J.; Schnieder, S.; Epps, J.; Quatieri, T.F. A Review of Depression and Suicide Risk Assessment Using Speech Analysis. *Speech Commun.* **2015**, *71*, 10–49. https://doi.org/10.1016/j.specom.2015.03.004.

21. Weeks, J.W.; Srivastav, A.; Howell, A.N.; Menatti, A.R. "Speaking More than Words": Classifying Men with Social Anxiety Disorder via Vocal Acoustic Analyses of Diagnostic Interviews. *J. Psychopathol. Behav.* **2016**, *38*, 30–41. https://doi.org/10.1007/s10862-015-9495-9.

22. Teferra, B.G.; Borwein, S.; DeSouza, D.D.; Rose, J. Screening for Generalized Anxiety Disorder from Acoustic and Linguistic Features of Impromptu Speech: Prediction Model Evaluation Study. *JMIR Form. Res.* **2022**, *6*, e39998. https://doi.org/10.2196/39998.

23. Goberman, A.M.; Hughes, S.; Haydock, T. Acoustic Characteristics of Public Speaking: Anxiety and Practice Effects. *Speech Commun.* **2011**, *53*, 867–876. https://doi.org/10.1016/j.specom.2011.02.005.

24. Albuquerque, L.; Valente, A.R.S.; Teixeira, A.; Figueiredo, D.; Sa-Couto, P.; Oliveira, C. Association between Acoustic Speech Features and Non-Severe Levels of Anxiety and Depression Symptoms across Lifespan. *PLoS ONE* **2021**, *16*, e0248842. https://doi.org/10.1371/journal.pone.0248842.

25. McGinnis, E.W.; Anderau, S.P.; Hruschak, J.; Gurchiek, R.D.; Lopez-Duran, N.L.; Fitzgerald, K.; Rosenblum, K.L.; Muzik, M.; McGinnis, R.S. Giving Voice to Vulnerable Children: Machine Learning Analysis of Speech Detects Anxiety and Depression in Early Childhood. *IEEE J. Biomed. Health Inform.* **2019**, *23*, 2294–2301. https://doi.org/10.1109/JBHI.2019.2913590.

26. Laukka, P.; Linnman, C.; Åhs, F.; Pissiota, A.; Frans, Ö.; Faria, V.; Michelgård, Å.; Appel, L.; Fredrikson, M.; Furmark, T. In a Nervous Voice: Acoustic Analysis and Perception of Anxiety in Social Phobics' Speech. *J. Nonverbal Behav.* **2008**, *32*, 195–214. https://doi.org/10.1007/s10919-008-0055-9.

27. Ben-David, B.M.; Multani, N.; Shakuf, V.; Rudzicz, F.; Van Lieshout, P.H.H.M. Prosody and Semantics Are Separate but Not Separable Channels in the Perception of Emotional Speech: Test for Rating of Emotions in Speech. *J. Speech Lang. Hear. Res.* **2016**, *59*, 72–89. https://doi.org/10.1044/2015_JSLHR-H-14-0323.

28. Olah, J.; Diederen, K.; Gibbs-Dean, T.; Kempton, M.J.; Dobson, R.; Spencer, T.; Cummins, N. Online Speech Assessment of the Psychotic Spectrum: Exploring the Relationship between Overlapping Acoustic Markers of Schizotypy, Depression and Anxiety. *Schizophr. Res.* **2023**, *259*, 11–19. https://doi.org/10.1016/j.schres.2023.03.044.

29. Demiris, G.; Oliver, D.P.; Washington, K.T.; Chadwick, C.; Voigt, J.D.; Brotherton, S.; Naylor, M.D. Examining Spoken Words and Acoustic Features of Therapy Sessions to Understand Family Caregivers' Anxiety and Quality of Life. *Int. J. Med. Inf.* **2022**, *160*, 104716. https://doi.org/10.1016/j.ijmedinf.2022.104716.

30. Sigmund, M. Influence of Psychological Stress on Formant Structure of Vowels. *Elektron. Elektrotech.* **2012**, *18*, 45–48. https://doi.org/10.5755/j01.eee.18.10.3059.

31. Banse, R.; Scherer, K.R. Acoustic Profiles in Vocal Emotion Expression. *J. Pers. Soc. Psychol.* **1996**, *70*, 614–636. https://doi.org/10.1037/0022-3514.70.3.614.

32. Hagenaars, M.A.; Van Minnen, A. The Effect of Fear on Paralinguistic Aspects of Speech in Patients with Panic Disorder with Agoraphobia. *J. Anxiety Disord.* **2005**, *19*, 521–537. https://doi.org/10.1016/j.janxdis.2004.04.008.

33. Scherer, K.R.; Grandjean, D.; Johnstone, T.; Klasmeyer, G.; Bänziger, T. Acoustic Correlates of Task Load and Stress. In Proceedings of the 7th International Conference on Spoken Language Processing, Denver, CO, USA, 16–20 September 2002; pp. 2017–2020.

34. Johannes, B.; Wittels, P.; Enne, R.; Eisinger, G.; Castro, C.A.; Thomas, J.L.; Gerzer, R. Non-Linear Function Model of Voice Pitch Dependency on Physical and Mental Load. *Eur. J. Appl. Physiol.* **2007**, *101*, 267–276. https://doi.org/10.1007/s00421-007-0496-6.

35. Van Puyvelde, M.; Neyt, X.; McGlone, F.; Pattyn, N. Voice Stress Analysis: A New Framework for Voice and Effort in Human Performance. *Front. Psychol.* **2018**, *9*, 1994. https://doi.org/10.3389/fpsyg.2018.01994.

36. Kirchhuebel, C. The Effects of Lombard Speech on Vowel Formant Measurements. In Proceedings of the São Paulo School of Advanced Studies in Speech Dynamics SPSASSD 2010 Accepted Papers, 7–11 June 2010; p. 38.

37. Mendoza, E.; Carballo, G. Acoustic Analysis of Induced Vocal Stress by Means of Cognitive Workload. *J. Voice* **1998**, *12*, 263–277. https://doi.org/10.1016/s0892-1997(98)80017-9.

38. Godin, K.W.; Hansen, J.H. Physical Task Stress and Speaker Variability in Voice Quality. *EURASIP J. Audio Speech Music Process.* **2015**, *2015*, 1–13.

39. Godin, K.W.; Hasan, T.; Hansen, J.H.L. Glottal Waveform Analysis of Physical Task Stress Speech. In Proceedings of the 13th Annual Conference of the International Speech Communication Association, Portland, OR, USA, 9–13 September 2012.

40. Koblick, H. Effects of Simultaneous Exercise and Speech Tasks on the Perception of Effort and Vocal Measures in Aerobic Instructors. Ph.D. Dissertation, University of Central Florida, Orlando, FL, USA, 2004.

41. Baird, A.; Cummins, N.; Schnieder, S.; Krajewski, J.; Schuller, B.W. An Evaluation of the Effect of Anxiety on Speech—Computational Prediction of Anxiety from Sustained Vowels. In Proceedings of the Interspeech 2020, Shanghai, China, 25–29 October 2020.

42. Borsboom, D.; Deserno, M.K.; Rhemtulla, M.; Epskamp, S.; Fried, E.I.; McNally, R.J.; Robinaugh, D.J.; Perugini, M.; Dalege, J.; Costantini, G.; et al. Network Analysis of Multivariate Data in Psychological Science. *Nat. Rev. Methods Primers* **2021**, *1*, 58.

43. Liu, C.; Ren, L.; Rotaru, K.; Liu, X.; Li, K.; Yang, W.; Li, Y.; Wei, X.; Yücel, M.; Albertella, L. Bridging the Links between Big Five Personality Traits and Problematic Smartphone Use: A Network Analysis. *J. Behav. Addict.* **2023**, *12*, 128–136. https://doi.org/10.1556/2006.2022.00093.

44. van Borkulo, C.D.; van Bork, R.; Boschloo, L.; Kossakowski, J.J.; Tio, P.; Schoevers, R.A.; Borsboom, D.; Waldorp, L.J. Comparing Network Structures on Three Aspects: A Permutation Test. *Psychol. Methods* **2023**, *28*, 1273–1285. https://doi.org/10.1037/met0000476.

45. Epskamp, S.; Fried, E.I. A Tutorial on Regularized Partial Correlation Networks. *Psychol. Methods* **2018**, *23*, 617–634. https://doi.org/10.1037/met0000167.

46. Endler, N.S.; Kocovski, N.L. State and Trait Anxiety Revisited. *J. Anxiety Disord.* **2001**, *15*, 231–245. https://doi.org/10.1016/S0887-6185(01)00060-3.

47. Julian, L.J. Measures of Anxiety: State-Trait Anxiety Inventory (STAI), Beck Anxiety Inventory (BAI), and Hospital Anxiety and Depression Scale-Anxiety (HADS-a). *Arthritis Care Res.* **2011**, *63*, S467–S472. https://doi.org/10.1002/acr.20561.

48. Liu, M.; Zhang, X. An Investigation of Chinese University Students' Foreign Language Anxiety and English Learning Motivation. *Engl. Linguist. Res.* **2013**, *2*, 1–13. https://doi.org/10.5430/elr.v2n1p1.

49. Hewitt, E.; Stephenson, J. Foreign Language Anxiety and Oral Exam Performance: A Replication of Phillips's MLJ Study. *Mod. Lang. J.* **2012**, *96*, 170–189. https://doi.org/10.1111/j.1540-4781.2011.01174.x.

50. Xu, Y.; Xie, Z. Exploring the Predictors of Foreign Language Anxiety: The Roles of Language Proficiency, Language Exposure, and Cognitive Control. *Front. Psychiatry* **2024**, *15*, 1492701. https://doi.org/10.3389/fpsyt.2024.1492701.

51. Ran, C.; Wang, Y.; Zhu, W. Comparison of Foreign Language Anxiety Based on Four Language Skills in Chinese College Students. *BMC Psychiatry* **2022**, *22*, 558. https://doi.org/10.1186/s12888-022-04201-w.

52. Qian, D.D.; Cumming, A. Researching English Language Assessment in China: Focusing on High-Stakes Testing. *Lang. Assess. Q.* **2017**, *14*, 97–100. https://doi.org/10.1080/15434303.2017.1295969.

53. Huang, H.-T.D.; Hung, S.-T.A. Comparing the Effects of Test Anxiety on Independent and Integrated Speaking Test Performance. *TESOL Quart.* **2013**, *47*, 244–269. https://doi.org/10.1002/tesq.69.

54. Titze, I.R.; Martin, D.W. *Principles of Voice Production*; National Center for Voice and Speech: Denver, CO, USA, 1998.

55. Kendall, T. *Speech Rate, Pause, and Sociolinguistic Variation: Studies in Corpus Sociophonetics*; Springer: Berlin/Heidelberg, Germany, 2013; ISBN 978-1-349-32095-0.

56. Jones, P.; Ma, R.; McNally, R. Bridge Centrality: A Network Approach to Understanding Comorbidity. *Multivar. Behav. Res.* **2019**, *56*, 353–367.

57. Isvoranu, A.; Epskamp, S. Which Estimation Method to Choose in Network Psychometrics? Deriving Guidelines for Applied Researchers. *Psychol. Methods* **2023**, *28*, 925–946.

58. Friedman, J.; Hastie, T.; Tibshirani, R. Sparse Inverse Covariance Estimation with the Graphical Lasso. *Biostatistics* **2008**, *9*, 432–441.

59. Foygel, R.; Drton, M. Extended Bayesian Information Criteria for Gaussian Graphical Models. In Proceedings of the Advances in Neural Information Processing Systems, Vancouver, BC, Canada, 6–9 December 2010; Volume 23, pp. 2020–2028.

60. Fruchterman, T.M.J.; Reingold, E.M. Graph Drawing by Force-Directed Placement. *Softw. Pract. Exper.* **1991**, *21*, 1129–1164. https://doi.org/10.1002/spe.4380211102.

61. Epskamp, S.; Cramer, A.O.J.; Waldorp, L.J.; Schmittmann, V.D.; Borsboom, D. Qgraph: Network Visualizations of Relationships in Psychometric Data. *J. Stat. Softw.* **2012**, *48*, 1–18. https://doi.org/10.18637/jss.v048.i04.

62. Robinaugh, D.; Millner, A.; McNally, R. Identifying Highly Influential Nodes in the Complicated Grief Network. *J. Abnorm. Psychol.* **2016**, *125*, 747–757.

63. Ren, L.; Wang, Y.; Wu, L.; Wei, Z.; Cui, L.-B.; Wei, X.; Hu, X.; Peng, J.; Jin, Y.; Li, F.; et al. Network Structure of Depression and Anxiety Symptoms in Chinese Female Nursing Students. *BMC Psychiatry* **2021**, *21*, 279. https://doi.org/10.1186/s12888-021-03276-1.

64. Epskamp, S.; Borsboom, D.; Fried, E.I. Estimating Psychological Networks and Their Accuracy: A Tutorial Paper. *Behav. Res. Methods* **2018**, *50*, 195–212. https://doi.org/10.3758/s13428-017-0862-1.

65.   Wei, X.; An, F.; Liu, C.; Li, K.; Wu, L.; Ren, L.; Liu, X. Escaping Negative Moods and Concentration Problems Play Bridge Roles in the Symptom Network of Problematic Smartphone Use and Depression. *Front. Public. Health* **2023**, *10*, 981136. https://doi.org/10.3389/fpubh.2022.981136.

66.   Teixeira, J.P.; Oliveira, C.; Lopes, C. Vocal Acoustic Analysis—Jitter, Shimmer and HNR Parameters. *Procedia Technol.* **2013**, *9*, 1112–1122. https://doi.org/10.1016/j.protcy.2013.12.124.

67.   Postma-Nilsenová, M.; Holt, E.; Heyn, L.; Groeneveld, K.; Finset, A. A Case Study of Vocal Features Associated with Galvanic Skin Response to Stressors in a Clinical Interaction. *Patient Educ. Couns.* **2016**, *99*, 1349–1354. https://doi.org/10.1016/j.pec.2016.03.006.

68.   Brenner, M.; Shipp, T.; Doherty, E.T.; Morrissey, P. Voice Measures of Psychological Stress: Laboratory and Field Data. In *Vocal Fold Physiology, Biomechanics, Acoustics, and Phonatory Control*; Titze, J.R., Scherer, R.C., Eds.; The Denver Center for the Performing Arts: Denver, CO, USA, 1985; pp. 239–248.

69.   Brenner, M.; Shipp, T. Voice Analysis under Stress. *J. Phon.* **1988**, *16*, 477–482.

70.   Yao, X.; Chen, B.; Yoshimura, H. Dynamic Features of Vocal Folds Based on Speech Production Model for Detection of Stressed Speech. In Proceedings of the International Conference Progress in Informatics and Computing (pic), Shanghai, China, 23–25 December 2016; IEEE: Shanghai, China, 2016; pp. 268–272.

71.   Kreiman, J.; Sidtis, D. *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*; Wiley-Blackwell: Hoboken, NJ, USA, 2011.

72.   Schröder, M.; Cowie, R.; Douglas-Cowie, E.; Westerdijk, M.; Gielen, S.C.A.M. Acoustic Correlates of Emotion Dimensions in View of Speech Synthesis. In Proceedings of the Interspeech, Aalborg, Denmark, 3–7 September 2001.

73.   Roy, N.; Bless, D.M. Personality Traits and Psychological Factors in Voice Pathology: A Foundation for Future Research. *J. Speech Lang. Hear. Res.* **2000**, *43*, 737–748. https://doi.org/10.1044/jslhr.4303.737.

74.   Scherer, K.R. Vocal Affect Expression: A Review and a Model for Future Research. *Psychol. Bull.* **1986**, *99*, 143–165.

75.   Murray, I.R.; Baber, C.; South, A. Towards a Definition and Working Model of Stress and Its Effects on Speech. *Speech Commun.* **1996**, *20*, 3–12. https://doi.org/10.1016/S0167-6393(96)00040-4.

76.   Scherer, K.R. The Dynamic Architecture of Emotion: Evidence for the Component Process Model. *Cogn. Emot.* **2009**, *23*, 1307–1351. https://doi.org/10.1080/02699930902928969.

77.   Lindblom, B. Explaining Phonetic Variation: A Sketch of the H&H Theory. In *Speech Production and Speech Modelling*; Hardcastle, W.J., Marchal, A., Eds.; Springer Netherlands: Dordrecht, The Netherlands, 1990; pp. 403–439, ISBN 978-94-009-2037-8.

78.   Xu, C.; Zhu, L.; Wang, Y.; Wang, R.; Liu, C. The Effect of Foreign Language Speaking Anxiety on the Quality of Verbal Interaction: A fNIRS-Based Hyperscanning Study. *Acta Psychol. Sin.* **2023**, *55*, 1949. https://doi.org/10.3724/SP.J.1041.2023.01949.

79.   Fakes, J.; Bond, Z.S. Non-Native English Speakers' Stress Patterns in Words and Sentences. *Human Commun. Canada/Communication Humaine Canada* **1986**,*10*, 5–10.

80.   Kallio, H.; Kaźmierski, K. Reduction of Unstressed English Vowels by EFL Speakers with Different Language Backgrounds. In Proceedings of the 5th International Symposium on Applied Phonetics (ISAPh 2024), Tartu, Estonia, 30 September–2 October 2024,; ISCA: Singapore, 2024; pp. 38–42.

81.   Mora, J.C.; Mora-Plaza, I.; Bermejo Miranda, G. Speaking Anxiety and Task Complexity Effects on Second Language Speech. *Int. J. Appl. Linguist.* **2024**, *34*, 292–315. https://doi.org/10.1111/ijal.12494.

82.   Szyszka, M.; Peltonen, P.; Lintunen, P. Unravelling the Relationship between Language Anxiety and Foreign Language Speech Fluency in a Monologue Production. *J. Multiling. Multicult. Dev.* **2024**, 1–15. https://doi.org/10.1080/01434632.2024.2387149.

83.   Flege, J. Second Language Speech Learning: Theory, Findings and Problems. In *Speech Perception and Linguislic Experience: Issues in Cross-Language Research.*; York Press: Baltimore, MD, USA, 1995; pp. 229–273.

84.   Guion, S.G.; Flege, J.E.; Akahane-Yamada, R.; Pruitt, J.C. An Investigation of Current Models of Second Language Speech Perception: The Case of Japanese Adults' Perception of English Consonants. *J. Acoust. Soc. Am.* **2000**, *107*, 2711–2724. https://doi.org/10.1121/1.428657.