BMC Psychology

RESEARCH

Open Access

Cross-modal co-occurrence analysis of nonverbal behavior and content words, vocal emotion and prosody in individuals with subclinical depression

Liusheng Wang[1], Fuying Wu[2], Haiyan Zhang[3*] and Dongfang Lin[4]

Abstract

Background Most related research focuses on a single variable of verbal and nonverbal behaviors independently without considering their associations. Therefore, it is important to understand subclinical depression in the entire population.

Aims This study investigated the cross-modal co-occurrence of nonverbal behavior with vocal emotions, prosody, and content words in individuals with subclinical depression.

Methods A total of 70 participants assigned to the subclinical depression and control groups participated in structured interviews. Elan software was used to layer, transcribe, and annotate materials. A support vector machine was used to confirm the two models.

Results Cross-modal co-occurrence analysis revealed that the subclinical depression group mainly exhibited strong relationships between the nonverbal behavior "holding hands" and the words including "conflict," "hope" and "suicide," while the control group exhibited strong relationship between the nonverbal behavior "holding hands" and the content words including "happy," "despair" and "stress," and strong relationships of more nonverbal behaviors with more positive and negative words. The "pause" and "hesitation" of prosody were strongly associated nodes with the subclinical depression group, while "pause" and "delight" (vocal emotion) were strongly associated nodes with the control group. The accuracy rates of the two models through support vector machine were high and could be confirmed.

Conclusions The results of the cross-modal co-occurrence analysis revealed negative thoughts and moods of individuals with subclinical depression, whose nonverbal behavior was closely connected with verbal factors.

Keywords Subclinical depression, Nonverbal behavior, Vocal emotion and prosody, Cross-modal co-occurrence analysis, Content word, Support vector machine

*Correspondence:
Haiyan Zhang
zhanghaiyan@jmu.edu.cn
Full list of author information is available at the end of the article

Introduction

Interpersonal interaction entails the reciprocal exchange of multimodal signals. Specifically, verbal and nonverbal behaviors form the foundation of interpersonal language. The signaling functions of nonverbal behavior, along with the intricate layers of signals in face-to-face interpersonal communication, pose significant semantic and temporal integration challenges [1]. Within the realm of psychology, scholars have traditionally directed their attention towards the notion of a stable personality trait known as "ventilatory personality", where individuals' respiratory patterns exhibit enduring consistency over time, remaining stable and unchanging even across days [2]. The processing of linguistic and paralinguistic information intertwines as listeners decode the voices of speakers [3]. The brain's response to words is notably influenced by the volume of information conveyed through multimodal cues, underscoring the reliance of language comprehension on both verbal and nonverbal cues. The interplay between multimodal cues is dynamic, with the impact of each cue evolving based on informational input from other cues [4].

Individuals experiencing depression commonly display anhedonia, distorted self-perception, lack of motivation, and physical symptoms [5]. These symptoms are linked to a negative cognitive bias [6, 7], inhibitory dysfunction, challenges in processing negative stimuli [8], and a strong negative interpretation bias towards ambiguous information [9]. Researchers have started to investigate the significance of various nonverbal behaviors in depressed individuals, encompassing somatic, postural, facial, and phonological characteristics. A meta-analysis revealed that individuals recovering from major depressive disorder exhibit poorer cognitive performance in areas such as attention, working memory, and long-term memory compared to healthy individuals, with performance declining further in cases of recurrent depression [10].

Compared with healthy individuals, depressed individuals differ in verbal and facial visual information aspects [11]. Individuals with depression use the first-person singular pronouns more often on social media [12], and when writing essays [13], and poems [14], revealing a strong relationship between the use of first-person singular pronouns and depression [15–17]. People with depression also use a greater proportion of past-focused words (e.g., "before," "done," and the past tense) and sad emotional words (e.g., "sadness," "crying") [18]. There exists the difference between depressed and healthy persons in processing visual information indicated by the 2D (two-dimensional) and 3D (three-dimensional) information [19].

Similarly, there exist some differences in postural control, motor activity, and body morphology between depressed individuals and healthy individuals. Patients with depression exhibit the reduced vertical head movement and a more slumped posture [20]. In addition to postural differences, depressed individuals are more likely to exhibit a hunchback, forward head posture, and rounded shoulders, and depression is significantly associated with spine abnormalities [21]. Depressed individuals exhibit motor activity more frequently at night, as well as higher frequencies and longer durations of self-touching, less eye contact, increased or decreased crying, fewer smiles, fewer eyebrow movements, fewer types of non-specific gaze fixations, more look-downward, and more gestures [11]. Body dysmorphia and deformity are risk factors for depression. A model using human joint data can distinguish between depressed and healthy individuals with high accuracy [22].

Individuals with depression exhibit specific vocal characteristics. Depressive states can be predicted by analyzing sounds, images, and semantic content [23], and vocal features can be used to effectively predict depression [24]. In terms of acoustic characteristics, depressed people have lower pitch variation, longer pauses, slower speech speed, and weaker lexical stress [11]. Voice abnormalities in patients with depression include cross-contextual stability, and potential behavioral indicators of depression used in voice recognition include loudness, MFCC5, MFCC7, jitter and cepstral peak prominence-smoothed (CPPS) [25, 26]. Mel frequency cepstral coefficients (MFCCs) refer to a set of features just like chroma or spectral, developed at MIT in the late 1960s to analyze seismic audio echoes and model human voice characteristics. Patients with major depressive disorders have less expressive prosody in their voices, which is likely to be accompanied by right hemisphere dysfunction [27]. Acoustic signatures are potential biomarkers of depression.

The studies mentioned above show that nonverbal behaviors such as posture and facial features differ between individuals with depression and healthy individuals. However, research on nonverbal expressions of emotions has mostly relied on facial expressions and overlooked the emotional expressions of the entire body. Additionally, measurements of nonverbal behavior in clinical populations lack ecological validity. Most related research has analyzed verbal and nonverbal behaviors independently without considering their associations. To enhance the understanding of the relationships between behavior, language, emotions, and cognitive components in interpersonal interactions [23], it is important to improve the ecological validity of experiments and use objective indicators. Therefore, to improve the ecological validity of the research conclusions, this study used a multimodal analysis method to investigate the associations between nonverbal behaviors (such as head posture, facial expressions, hand movements, body posture, and leg movements) and vocal emotions and prosody in individuals with subclinical depression.

The purpose of this study was to investigate the cross-modal co-occurrence of nonverbal behavior with vocal emotions, prosody, and content words in individuals with subclinical depression. Individuals with subclinical depression are expected to (1) exhibit a higher co-occurrence of nonverbal behaviors and content words compared to healthy individuals, and (2) show a higher co-occurrence of nonverbal behaviors with vocal emotions and prosody compared to healthy individuals.

Method

Participants

All the participants were recruited from universities and colleges. A total of 2849 college students volunteered to participate in the online survey. The Beck Depression Inventory (BDI-II-C) and Patient Health Questionnaire-9 (PHQ-9) were used as screening tools. Participants of subclinical depression are in a depressive state that did not meet the symptom and course criteria of MDD in DSM-IV [28, 29]. After completing the BDI-II-C and PHQ-9, 47 college students were assigned to the subclinical depression group and the control group comprised 23 college students. The participants had a mean age of 19.69 years (SD = 1.27) years. All the participants provided written informed consent and received compensation for their participation.

Tools

The Beck Depression Inventory (BDI-II-C) is a widely used self-report questionnaire for measuring depression in adults [5]. A Chinese version has been developed [30]. The scale of BDI-II-C has a total of 21 items, with a total score ranging from 0 to 63 points, with a larger total score indicating more severe depression. A total score is greater than or equal to 14 for mild depression and above. The National Health Commission of China recommends that medical and health institutions use the Patient Health Questionnaire (PHQ-9) to screen for and assess the severity of depression [31–33].The scale of PHQ-9 has a total of 9 items with a total score ranging from 0 to 27 points. A total score greater than or equal to 5 is mild depression and above. The higher the score, the more severe the depression.

The BDI scores in the subclinical and control group were $M1 = 26.40$, $SD1 = 8.84$ and $M2 = 2.70$, $SD21 = 2.77$, respectively. An independent-sample t-test showed that the BDI score of the subclinical group was significantly higher than that of the control group [$t\ (68) = 16.79$, $p < 0.001$]. The PHQ-9 scores of the subclinical and control groups were $M1 = 13.68$, $SD1 = 4.76$; and $M2 = 1.57$, $SD2 = 1.34$, respectively. The PHQ-9 score of the subclinical group was significantly higher than that of the control group $t\ (68) = 16.20$, $p < 0.001$.

Outline of the interviews. The interviews were a self-exploration process. Considering Beck's depressive cognitive triad, Bronfenbrenner's ecological systems theory, and Mead's dual feedback loop model, the interview outline covered the self and emotions, family relationships, social relationships, social events, and the future of life. The content of the interview outline received evaluations and feedback from hospital professional clinical psychiatrists, psychological counselors, psychology professors, and experts (see Supplement 1).

Procedure

First, the participants were screened using the two scales of the online survey. Second, the participants were interviewed individually according to the topics of the interview outline, and their entire bodies were recorded during the interviews. The interview topics revolved around the self and emotions, family relationships, social relationships, social events, and the future. Third, the data of videos were conducted including segmentation and annotation in a tier-by-tier manner by Elan software. Relevant behaviors were identified and labeled according to the annotation tiers. In addition, the speech in the video was transcribed into text. The total length of the videos was 1134.08 min. Finally, a support vector machine was used to confirm the two models.

Data analysis

Data analysis by Elan software

Elan software is used to annotate video and audio files [34]. Annotation describe certain features of the video and audio files using sentences, vocabulary, etc.. The videos were segmented and annotated in a tier-by-tier manner. Relevant behaviors were identified and labeled according to the annotation tiers. In the current study, the annotation tiers in the Elan software included head posture, facial expressions, hand movements, body posture, leg movements, vocal emotions, and prosody. Annotations can be divided into different tiers according to the attributes of the described features. These tiers are time-locked. So, the data through Elan software include the frequency and the duration of variable, which is helpful to analyze nonverbal behaviors or co-occurrence networks.

Co-occurrence analysis

In order to better understand the relationships between nonverbal behaviors and verbal cues in subclinical depression, co-occurrence analysis was conducted.

Co-occurrence analysis generates a co-occurrence network and matrix of co-occurrence probability coefficients based on the annotation frequency and annotation duration [35]. A matrix of co-occurrence probability coefficients can indicate relationships among subcategories in different modalities based on some formulas proposed by Wang et al. [36].

Co-occurrence analysis was based on two indicators, annotation frequency and annotation duration, in Elan software. The annotation frequency is the number of occurrences divided by the observation period and the annotation duration is the duration of the annotations divided by the observation period. Duration and frequency are both important and good indices for co-occurrence analysis. The degree of co-occurrence is related to duration and frequency, similar to the two dimensions of a coordinate system that are associated with different characteristics. If they are combined into an intuitive index, the specific weights of these two dimensions on the degree of co-occurrence cannot be scientifically determined. Therefore, in this study, co-occurrence analysis of these two indicators was conducted separately. The coefficients of co-occurrence probability in the co-occurrence analysis indicated associations among different factors in a specific group. Coefficients from the same matrix can be compared; however, coefficients from different matrices cannot be directly compared with or without parameter tests.

Support vector machine

A support vector machine (SVM) was used to confirm the two models. The model 1 was a co-occurrence model of nonverbal behavior and content words, and the model 2 was a co-occurrence model of nonverbal behavior with vocal emotion and prosody. A support vector machine (SVM) is a supervised machine learning algorithm that classifies data by finding an optimal line that maximizes the distance between each class in an N-dimensional space. Typically, the dataset is divided into two sets: training set and test set. A training set is used to train the model. A test set was primarily used to evaluate the generalization performance of the model. In this study, SVM was applied to confirm the models presented in accuracy rate and learning curve.

Results

Cross-modal co-occurrence analysis of nonverbal behavior and content words

This study focused only on content words and not function words. A total of 83 content words were included in the cross-modal association analysis (see Table 1). In terms of annotation frequency and annotation duration, the strongest associations between content words and nonverbal behavior in the subclinical depression group were between HH (holding hands) and Cfl (conflict), Hope, and Suic (suicide) (Table 2, Supplement 2 Table 1, Supplement 2 Table 2, and Abbreviations).

In terms of annotation frequency, the strongest associations in the control group were as follows: Cfl (conflict) with HH (holding hands), LAR (look around), TT (touching things), PFT (putting feet together), HN (head nod), SM (smile), OFOB (one foot in front and one behind), SB (shake body), LEA (lean against), PS (pause), and TIH (tilting head); Hope with PFT (feet together), HH (holding hands), TT (touching things), SM (smile), LAR (look around), SB (shake body), and LS (look straight); Suic (suicide) with LAR (look around), TT (touching things), PFT (putting feet together), and SB (shake body); Happ (happy) with SM (smile), DE (delight), TT (touching things), SB (shake body), HH (holding hands), LS (look straight), RTS (raising the tone suddenly), and SWL (swing legs); Cfor (comfortable) with TT (touching things) and PFT (putting feet together); Despair with HH (holding hands), SM (smile), and PFT (putting feet together); Boring with SM (smile) and PFT (putting feet together); Cfuse (confused) with TT (touching things); Unple (unpleasant) with PFT (putting feet together); and Stress with HH (holding hands) and LEA (lean against) (Table 3, Supplement 2 Table 3).

Table 1 The list of content words

Categories	Content words
Negative	Anger, Angry, Annoy, Afraid, Anxiety, Argue, Boring, Bothering, Conflict, Confused, Corrupted, Cry, Dark, Dead, Death, Decadence, Depression, Despair, Distressed, Dispirited, Downcast, Dreary, Escape, Exhausted, Fear, Flee, Frustrated, Frenzied, Grieved, Gloom, Grief, Hate, Impulsion, Indignant, Irritating, Joyous, Lacrimation, Loss, Numb, Oppressing sensation, Pathetic, Patient, Pain, Pessimism, Quarrel, Repression, Sad, Sensitive, Shiver, Slump, Sorrow, Sorriness, Somber, Stress, Suicide, Suffocative, Stimulated, Stubborn, Testiness, Troubled, Uninteresting, Unpleasant, Whiny, Weary
Positive	Comfort, Confidence, Delight, Friend, Future, Glad, Hope, Happy, Intimacy, Like, Lonely, Me, Mom, Mother, Nice, Optimism, Parents, Pleasant, Satisfaction

Table 2 The matrix of co-occurrence probability coefficients between nonverbal behaviors and content words based on the annotation frequency and duration in the subclinical group

Item	Index	Cfl (conflict)	Hope	Suic (suicide)
HH (holding hands)	Frequency	0.00013	0.00010	0.00013
	Duration	0.00004	0.00003	0.00004

Table 3 The matrix of co-occurrence probability coefficients indicating associations between nonverbal behaviors and content words based on the annotation frequency in control group

	Cfl	Hope	Suic	Happ	Cfor	Desp	Bor	Cfuse	Unple	Stres
HN	0.00017	0.00004	0.00000	0.00009	0.00004	0.00000	0.00000	0.00002	0.00002	0.00000
LS	0.00006	0.00011	0.00006	0.00013	0.00009	0.00004	0.00004	0.00009	0.00004	0.00002
TIH	0.00011	0.00002	0.00004	0.00004	0.00004	0.00002	0.00009	0.00000	0.00000	0.00004
HH	0.00045	0.00021	0.00006	0.00015	0.00006	0.00021	0.00009	0.00004	0.00002	0.00021
TT	0.00023	0.00017	0.00019	0.00021	0.00021	0.00004	0.00006	0.00011	0.00006	0.00002
SM	0.00017	0.00015	0.00006	0.00026	0.00006	0.00019	0.00019	0.00002	0.00002	0.00002
LAR	0.00028	0.00017	0.00021	0.00006	0.00006	0.00002	0.00002	0.00006	0.00000	0.00004
DE	0.00000	0.00000	0.00000	0.00021	0.00002	0.00006	0.00006	0.00000	0.00000	0.00000
RTS	0.00000	0.00002	0.00002	0.00013	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
PS	0.00013	0.00009	0.00002	0.00000	0.00000	0.00000	0.00000	0.00004	0.00002	0.00004
SB	0.00015	0.00011	0.00013	0.00019	0.00006	0.00002	0.00009	0.00004	0.00002	0.00002
LEA	0.00013	0.00002	0.00000	0.00000	0.00006	0.00002	0.00009	0.00002	0.00000	0.00013
SWL	0.00009	0.00009	0.00006	0.00011	0.00006	0.00002	0.00004	0.00000	0.00000	0.00002
OFOB	0.00017	0.00002	0.00004	0.00002	0.00002	0.00000	0.00002	0.00002	0.00000	0.00002
PFT	0.00028	0.00034	0.00017	0.00006	0.00013	0.00019	0.00013	0.00002	0.00011	0.00004

In terms of annotation duration, the strongest associations in the control group were as follows: Cfl (conflict) with HH (holding hands), LAR (look around), TT (touching things), PFT (putting feet together), SM (smile), PS (pause), OFOB (one foot in front and the other in back), SB(shake body), and HN (head nodding); Hope with PFT (putting feet together), TT (touching things), SM (smile), LAR (look around), and SB (shake body); Suic (suicide) with LAR (look around), SB (shake body), and TT (touching things); Happ (happy) with SM (smile), SB (shake body), DE (delight), HH (holding hands), TT (touching things), and LS (look straight); Cfor (comfort) with TT (touching things); Despair with HH (holding hands), SM (smile), and PFT (putting feet together); Bor (boring) with SM (smile); Unple (unpleasant) with PFT (putting feet together) (Table 4, Supplement 2 Table 4).

In short, the subclinical depression group exhibited a strong relationship between nonverbal behavior “holding hands” and content words, including “conflict”, “hope”, and “suicide”. The control group exhibited strong

Table 4 The matrix of co-occurrence probability coefficients indicating associations between nonverbal behaviors and content words based on the annotation duration in control group

	Cfl	Hope	Suic	Happ	Cfor	Desp	Bor	Cfuse	Unple	Stres
HN	0.00004	0.00001	0.00000	0.00001	0.00002	0.00000	0.00000	0.00001	0.00001	0.00000
LS	0.00002	0.00003	0.00001	0.00004	0.00002	0.00001	0.00001	0.00002	0.00002	0.00000
TIH	0.00003	0.00001	0.00001	0.00001	0.00000	0.00001	0.00001	0.00000	0.00000	0.00000
HW	0.00003	0.00003	0.00002	0.00003	0.00000	0.00000	0.00002	0.00000	0.00000	0.00001
HH	0.00012	0.00003	0.00002	0.00005	0.00002	0.00006	0.00002	0.00001	0.00001	0.00003
TT	0.00007	0.00005	0.00004	0.00005	0.00004	0.00001	0.00002	0.00003	0.00003	0.00000
SM	0.00005	0.00004	0.00001	0.00007	0.00002	0.00005	0.00005	0.00001	0.00001	0.00001
LAR	0.00008	0.00004	0.00005	0.00002	0.00001	0.00001	0.00000	0.00002	0.00000	0.00000
DE	0.00000	0.00000	0.00000	0.00006	0.00000	0.00002	0.00002	0.00000	0.00000	0.00000
RTS	0.00000	0.00001	0.00000	0.00003	0.00000	0.00000	0.00000	0.00000	0.00000	0.00000
PS	0.00003	0.00002	0.00001	0.00000	0.00000	0.00000	0.00000	0.00001	0.00001	0.00001
SB	0.00004	0.00004	0.00004	0.00006	0.00002	0.00001	0.00002	0.00002	0.00001	0.00001
LEA	0.00003	0.00000	0.00000	0.00000	0.00001	0.00000	0.00001	0.00001	0.00000	0.00001
SWL	0.00003	0.00002	0.00002	0.00003	0.00001	0.00001	0.00000	0.00000	0.00000	0.00001
OFOB	0.00005	0.00000	0.00002	0.00001	0.00000	0.00000	0.00001	0.00001	0.00000	0.00000
PFT	0.00007	0.00007	0.00003	0.00002	0.00002	0.00004	0.00003	0.00001	0.00004	0.00001

relationships between "holding hands" and the words including "conflict," "hope," "happy," "despair," and "stress," as well as strong relationships of more nonverbal behaviors with additional positive and negative words, and a strong association of the word "happy" with some nonverbal behaviors such as "smile" (facial expression), "delight" (vocal emotion), "touching things" (hand movement), "shake bode" (body posture).

Two methods of SVM and Random Forest were used for verification. The results showed that SVM was used for training due to its high efficiency in processing high-latitude data. SVMs can be used to make them faster and more accurate in this study. SVM was applied to confirm the models. The characteristics of the subclinical depression group were taken as the inclusion conditions, including the high co-occurrence relationship between the nonverbal behavior "holding hands" and the content words including "conflict", "hope", and "suicide". The characteristic of the control group was taken as the excluding conditions, including high co-occurrence relationships between "holding hands" and the words including "happy," "despair," and "stress," as well as a strong association of the word "happy" with some nonverbal behaviors such as "smile" (facial expression), "delight" (vocal emotion), "touching things" (hand movement), "shake bode" (body posture).

The SVM analysis showed that the accuracy rate was 76% for both the frequency and duration of the annotation. The training score curve first decreased and then increased gradually (Fig. 1). This means that the model may initially have some overfitting on the training set, that is, the performance on the training data is relatively high, but the model does not fully generalize to unseen data. Then, as the size of the training samples increased, the degree of overfitting gradually decreased, resulting in a flat increase in the training score curve. The cross-validation score curve slowly increased and then flattened (Fig. 1). This indicates that the performance of the model on the validation dataset gradually improved, with no significant improvement, even after adding more training data. This may indicate that the model has learned most of the features of the data and can generalize well to unseen data. Considering these two cases, the SVM learning curve of the model was good.

Cross-modal co-occurrence analysis of nonverbal behavior with vocal emotion and prosody

This section mainly refers to the relationships among head posture, hand movements, facial expressions, body posture, leg movements, vocal emotions, and prosody. In terms of annotation frequency, the strongest associations in the subclinical depression group were as follows: HES (hesitation) with OL (open legs), LS (look straight), HH (holding hands), LD (look down), and STR (straight); PS (pause) with OL (open legs), LS (look straight), HH (holding hands), STR (straight), and LA (look aside); RTS (raising the tone suddenly) with SWL (swing legs) and SB (shake body) (Table 5, Supplement 2 Table 5).

The strongest associations in the control group were as follows: DE (delight) with SM (smile), SB (shake body), TT (touching objects), HH (holding hands), LS (look straight), PFT (putting feet together), and PS (pause) with LEA (lean against), HH (holding hands), TT (touching things), OFOB (one foot in front and one behind), PFT (putting feet together), LS (look straight), LAR (look around), SB (shake body), TIH (tilting head), and TH (twisting head) (Table 6, Supplement 2 Table 6).

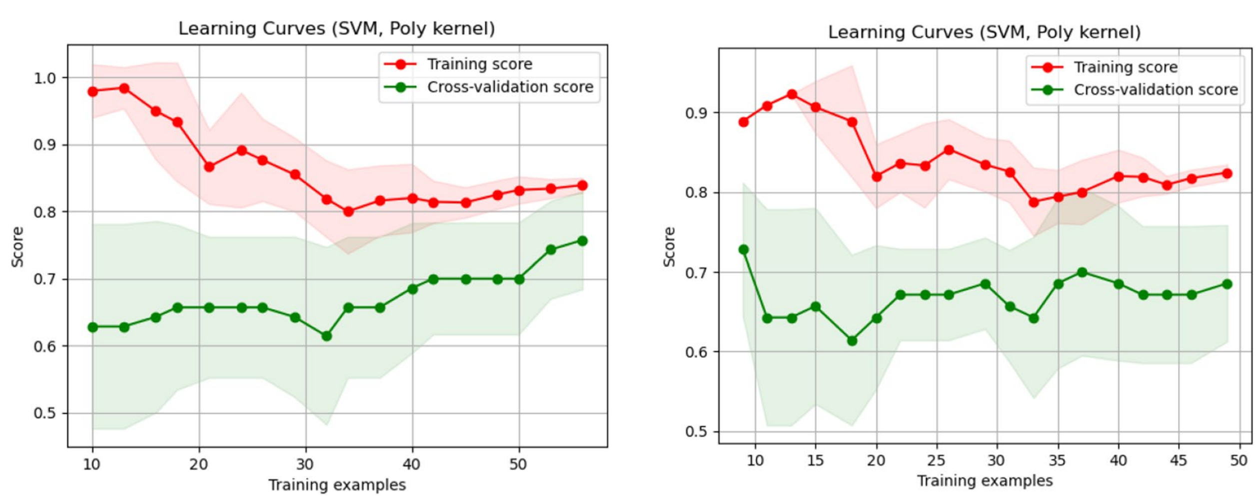

Fig. 1 The SVM learning curves involving the duration (left) and the frequency of annotation (right) for the model 1

Table 5 The matrix of co-occurrence probability coefficients indicating associations of nonverbal behaviors with vocal emotion and prosody based on the annotation frequency in subclinical group

	HES	RTS	PS
LS	0.0029	0.0013	0.0036
HH	0.0022	0.0008	0.0035
LA	0.0010	0.0010	0.0020
LD	0.0021	0.0003	0.0017
SB	0.0012	0.0023	0.0018
STR	0.0020	0.0012	0.0022
SWL	0.0012	0.0026	0.0017
OL	0.0036	0.0005	0.0048

Table 6 The matrix of co-occurrence probability coefficients indicating associations of nonverbal behaviors with vocal emotion and prosody based on the annotation frequency in control group

	DE	PS
TH	0.0004	0.0032
LS	0.0024	0.0048
TIH	0.0011	0.0038
HH	0.0031	0.0088
TT	0.0036	0.0084
SM	0.0053	0.0019
LAR	0.0004	0.0045
SB	0.0037	0.0043
LEA	0.0013	0.0130
OFOB	0.0013	0.0067
PFT	0.0022	0.0051

In terms of annotation duration, the strongest associations in the subclinical depression group were as follows: HES (hesitation) with OL (open legs), LS (look straight), STR (straight), and HH (holding hands); and PS (pause) with OL (open legs), SB (shaking body), HH (holding hands), LS (look straight), HW (head wagging), TIP (tiptoe), and STR (straight) (Table 7, Supplement 2 Table 7).

The strongest associations in the control group were as follows: DE (delight) with SM (smile), SB (shake body), HH (holding hands), TT (touching things), and LS (look straight); and PS (pause) with LEA (lean against), TT (touching things), HH (holding hands), OFOB (one foot in front and one behind), LS (look straight), LAR (look around), SB (shaking body), and PFT (putting feet together) (Table 8, Supplement 2 Table 8).

In short, "pause" (prosody) was strongly associated with "opening legs"(leg movement) and "holding hand" (hand movement), and, "hesitation" (prosody) was strongly associated with "opening legs"(leg movement) and "look straight" (head posture) in the subclinical depression group. While "pause" was strongly associated with "lean against"(body posture), "delight"(vocal emotion) was strongly associated with "smile" (facial expression), and "excited" (vocal emotion) was strongly associated with "putting feet together" (body posture) in the control group.

Table 7 The matrix of co-occurrence probability coefficients indicating associations of nonverbal behaviors with vocal emotion and prosody based on the annotation duration in subclinical group

	HES	PS
LS	0.0029	0.0037
HW	0.0007	0.0037
HH	0.0023	0.0042
SB	0.0013	0.0045
STR	0.0028	0.0028
TIP	0.0004	0.0031
OL	0.0042	0.0051

Table 8 The matrix of co-occurrence probability coefficients indicating associations of nonverbal behaviors with vocal emotion and prosody based on the annotation duration in control group

	DE	PS
LS	0.0020	0.0037
HH	0.0034	0.0051
TT	0.0027	0.0056
SM	0.0052	0.0012
LAR	0.0003	0.0033
SB	0.0038	0.0030
LEA	0.0004	0.0059
OFOB	0.0007	0.0038
PFT	0.0012	0.0030

SVM was applied to confirm the models. The characteristic of the subclinical depression group was taken as the including conditions, including "pause" (prosody) was high co-occurrence with "look straight," "holding hand" (hand movement), "straight," "opening legs"(leg movement), "shake body"; and "hesitation" (prosody) was strongly associated with "look straight" (head posture), "holding hand," "straight," "opening legs"(leg movement), "look down." The characteristic of the control group was taken as the excluding conditions, including high co-occurrence relationships between "pause" with "lean against," "delight" with "smile," and "excited" with "putting feet together." SVM analysis showed that the accuracy rate was 84% for the frequency of annotations and 81% for the duration of annotations. The training score curve first decreased and then increased gradually (Fig. 2). The training score drops first, probably because the model starts to learn the features and patterns of the data; however, over time, the model becomes more accurate, so the training score steadily improves. The cross-validation score curve first increases, then decreases, and then flattens out. The cross-validation score increases at the beginning, which indicates that the model's performance on the cross-validation data gradually improves but then declines, which may be due to the model overfitting the training data, resulting in a decrease in the performance of the cross-validation data. Finally, flattening indicates that the model has found an appropriate level of complexity to maintain a consistent performance across different validation sets. Considering these two cases, the SVM learning curve of the model conformed to a general pattern.

Discussion

In terms of the co-occurrence of nonverbal behaviors and content words, based on annotation frequency and annotation duration, the strongest associations in individuals with subclinical depression were for the behavior of "holding hands" with the words of "conflict," "hope," and "suicide." The associations between other nonverbal behaviors and other content words were very weak. However, in the control group, more nonverbal behaviors and content words co-occurred, indicating a strong association. The control group exhibited strong associations of "holding hands" with the words of "conflict," "hope," "happy" and "despair." In particular, the word "suicide" was relative strong association with "holding hand" in the subclinical group, while the word "happy" was relative strong association with "smile" in the control group. The strongest associations in the subclinical group were a subset of those observed in the control group. Therefore, the Hypothesis 1 of the study was supported. There exists the high co-occurrence of some nonverbal behaviors and some content words in individuals with subclinical depression, which is different from that of the control group.

Three characteristics were obtained from the analysis of the co-occurrence of nonverbal behaviors and content words. First, the two groups had different high co-occurrence network of word "suicide." There was a strong association between the word "suicide" and the behavior "holding hands" in the subclinical group, while the word "suicide" was not strongly associated with "holding hands" but rather with other more nonverbal behaviors such as "look around" in the control group. Here, "holding hands" refers not to the interaction with others' hands but to an individual's two-hand touching action. "Holding hand" reflects the person's ability to maintain self-control by using his/her the other hand to steady his body and, consequently, his mind, which also consistent with the study by Chen et al. [8]. The correlation between the term "suicide" and elevated self-control behaviors in subclinical persons suggests that the subject matter and lexicon surrounding "suicide" are particularly delicate and relevant to subclinical individuals. "Look around" reflects greater relaxation and lower self-control. So, people from different groups hold a different attitude reflected by a high co-occurrence network of the word "suicide."

Furthermore, the two groups had different aggregations and dispersions in the high co-occurrence network. The subclinical group was strongly associated with three words: "conflict", "hope", and "suicide". In contrast to the subclinical group, the control group had strong

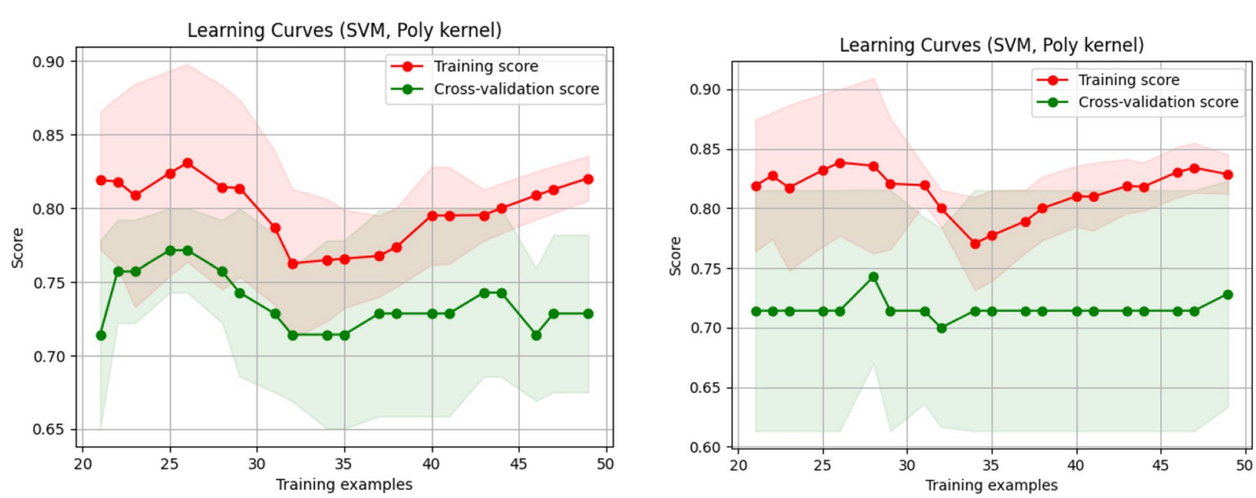

Fig. 2 The SVM learning curves involving the duration (left) and the frequency of annotation (right) for the model 2

relationships with more content words, such as "conflict," "hope," "happy," "despair," and "stress," with a variety of nonverbal behaviors. These words had both positive and negative meanings. Individuals in the subclinical group had more focused vocabulary and exerted greater self-control over nonverbal behavior, leading to stronger associations. Consequently, the control group exhibited a more diffuse resonance relationship between more words and more nonverbal behaviors, and healthy individuals did not exhibit a generally consistent negative cognitive bias or negative mood state. However, a stronger resonant link was observed between more narrowly focused speech and more tightly controlled nonverbal conduct in the subclinical group, particularly when it came to negative words and relatively highly regulated actions. The words of "conflict," "hope," and "suicide" had strong associations with individuals' emotional factors (in semantics) and the negative moods of individuals with subclinical depression. When individuals with subclinical depression use words such as "conflict," "hope," or "suicide," and these words are frequently accompanied by the nonverbal behavior "holding hand," indicating that verbal processing is strongly related to the control of nonverbal behavior. This is consistent with embodied cognition theory. Conceptual processing, such as the content words in this study, involves the partial reactivation or re-enactment of the feeling-action state that is experienced. Because of the presence of experiential information and the involvement of the emotional system in this process, conceptual processing shifts from the abstract to the concrete level [37]. Nonverbal behavior represents the physiological state or the affect-action state of the person or the emotional and emotional system, whereas the processing of words represents conceptual processing. Owing to the relationship between conceptual and emotional processing, the content words in this study were strongly related to specific nonverbal behaviors.

Third, the healthy people hold a resonance network with the word "happy." This mood was different from the mood of depressed individuals [6]. The control group exhibited strong resonance for the words "happy" and nonverbal behaviors like "smile"(facial expression), "delight" (vocal emotion), "touching things"(hand movement), "shake bode"(body posture), and "holding hand"(hand movement), which did not present in the subclinical group. Healthy people complement their speech with pleasant facial expressions, happy voices, and hand gestures to communicate with happy interior feelings when they use the word "happy" to describe themselves. Otherwhile, individuals of subclinical group communicated with others, without the word "happy" popping-out in the resonance network, lacking happy feelings.

Regarding the co-occurrence of nonverbal behaviors with vocal emotions and prosody, the Hypothesis 2 of the study was supported. There exists the high co-occurrence of some nonverbal behaviors and some vocal emotions and prosody in individuals with subclinical depression, which is different from that of the control group. the models with SVM were confirmed. The models arrived at high accuracy rates, and two points may be addressed. First, there was a difference in the nodes in the high co-occurrence network. "Hesitation" (prosody) was strongly associated with "opening legs"(leg movement), and "look straight" (head posture) in the subclinical depression group, while "delight"(vocal emotion) was strongly associated with "smile" (facial expression) in the control group. The associated with "hesitation" in the subclinical group reflects a lack of fluency in individual communication. It may be due to cognitive ambiguity, emotional hesitation, anxiety, or psychomotor retardation, which are typical of depressed patients [38]. The physical movements that match the "hesitation" of the prosody are more reflective of stillness. These relatively stillness movements of the body with high resonance are appropriately matched by the slowness (i.e., "hesitation") of prosody. The strongest association in the control group was "delight." The "delight" of vocal emotion was strongly associated with "smile" (facial expression) in the control group, which reflects a positive emotional state of individuals during interpersonal communication. This positive inner state is revealed by human voice. One of the views on the relationship between language and thinking is that language is a tool and material shell of thinking [39]. In many cases, people are hesitant and not fluent when their ideas are precisely transformed into external language, because they are not clear in cognition and thinking, and the logical level is not clear. The "delight" of vocal emotion and its high-resonance nonverbal behavior, that is, the joy (i.e., "delight") of vocal emotion, are accompanied by the movement of the body, depicting a picture of both form and spirit, overflowing with words, and acting with the whole body. The subclinical group lacked the vocal resonance node "delight," which meant that the individual vocal resonance in this group lacked positive emotional factors. In short, the slowness of prosody and the stillness of the action (or body), the joy of vocal emotion, and the movement of the action are interactively matched and are in line with the harmony and consistency of the body.

Secondly, the nonverbal behaviors associated with "pause" differed in two groups. "Pause" (prosody) was strongly associated with "opening legs" (leg movement), and "holding hands" (hand movement) in the subclinical depression group, while "pause" was strongly associated with "holding hands" (hand movement), "touching things" (hand movement) and "lean against" (body posture) in the control group. Pauses were associated with more varied nonverbal behaviors, such as body, head, hand, feet, and eye movements in the control group, whereas the subclinical group had fewer nodes with nonverbal behaviors and more controlled nonverbal behaviors. The consequences of the prosody "pause" of the two groups of individuals may be different. The nodes of nonverbal behaviors in the high co-occurrence network of subclinical group looked to be more rigid and more passive. The nodes of nonverbal behaviors in the high co-occurrence network of control group looked to be more flexible and more active. The difference between the two groups reflects the "principle of unity of speech, thought, affect and appearance" [40], not only in terms of the characteristics revealed in this study but also in the long-term physical and mental development of the two groups. These internal characteristics are revealed through speech and behavior in interpersonal communications.

Implication, limitation and future study

Cross-modal co-occurrence analysis of this study revealed strong relationships between some nonverbal behaviors and the words, vocal emotion and prosody in the individuals with subclinical depression. These associations were different from those of healthy people. These nonverbal behaviors included head posture, facial expressions, hand movements, body posture, and leg movements. These findings indicate a comprehensive way to recognize depressive or subclinical depression, not depending on the information from single mode. The information from cross-modality is ecological validity to analysis the depressive disorders. The negative thoughts and moods of individuals with subclinical depression could be represented by nonverbal behavior and verbal factors.

The findings of this study must be considered in light of the study's limitations. First, this study focused on the subclinical depressed people. Future studies should pay attention to clinical data and compare subclinical with clinical patients. Second, this study could not analyze the acoustic parameters of the sound. Future studies should attempt to examine the acoustic information of speech of subclinical depressed individuals. This requires more rigorous recording studios and more sophisticated recording equipment for sound acquisition. Third, people from different backgrounds and different cultures may have different understanding to the behaviors in the same situation. This study did not focus on the cultural factors which might influence the observed behaviors. Future studies can address the effects of cultural factors on the observed behaviors involving the subclinical depression.

Abbreviations

Head posture

Draw back	DB
Head drop	HD
Head nod	HN
Head shaking	HS
Head wagging	HW
Look straight	LS
Raising head	RH
Squint	SQ
Throw the head	TTH
Tilt head	TIH
Torsion head	TH

Hand movements

Beating and slapping	BS
Clapping	CL
Combing hair	CH
Covering mouth	COM
Draw back hands	DBH
Drooping hands	DH
Fingers opening and closing	FOC
Gesticulate	GE
Hand flat	HFL
Hand trembling	HT
Hold the fist in the other hand	HF
Holding hands	HH
Horizontal pointing	HP
Hugging	HUG
Making a fist	MF
Ok	OK
Palms opening and closing	POC
Picking at the hands	PAH
Pointing ahead	PA
Pointing oneself	PO
Putting hands in pocket	PHP
Raising hands	RAH
Rubbing hands	RUB
Scratch	SC
Spreading hands	SH
Swing hands	SWH
Thumb	THU
Touching chest	TC
Touching ear	TE
Touching eyes	TEY
Touching face	TF
Touching hands or wrist	THW
Touching jaw	TJ
Touching leg or knee	TLK
Touching neck	TNE
Touching nose	TN
Touching things	TT
Touching waist	TW
Vertical pointing	VP
Wave	WA

Body postures

Hunchback	HU
Lean against	LEA
Lean forward	LF
Shake body	SB
Shrug shoulders	SS
Straighten	STR
Tilting forward	TIF

Facial expressions

Bite lips	BLI
Blink	BL
Closing eyes	CE
Closing mouth	CLM

Evasive eye contact	EEC
Extend tongue	ET
Forced smile	FS
Frown	FR
Laugh	LAU
Lick lips	LL
Look around	LAR
Look aside	LA
Look down	LD
Look up	LU
Open mouth	OM
Pouting	POU
Puckering lips	PL
Query	QU
Raise eyebrow	RE
Screw up eyes	SUE
Shed tears	SHT
Smile	SM
Sneer	SN
SorrowFace	SOF
Stare blankly	STB
Staring	ST
Swallow	SW
Twitching mouth	TM
Leg movements	
Crossing feet	CRF
Crossing legs	CRL
Lifting the feet	LTF
One in front and the other in back	OFOB
Opening legs	OL
Putting feet together	PFT
Retracting legs	RL
Rubbing floor	RF
Stamp	STA
Stretch legs	STL
Swing legs	SWL
Tiptoe	TIP
Vocal emotion	
Admiring	AD
Anger	AN
Anxiety	ANX
Ask rhetorically	AR
Aversion	AV
Bitter and astingent	BA
Confused	CO
Delight	DE
Depressed	DEP
Helpless	HEL
Hesitation	HES
Excited	EX
Sorrow	SO
Prosody	
Cough	COU
Drawl	DR
Emphaticalness	EM
Intermittent sound	IS
Lowering the tone gradually	LTG
Lowering the tone suddenly	LTS
Pause	PS
Raising the tone	RTT
Raising the tone suddenly	RTS
Repetition	REP
Sigh	SIG
Speak faster	SPF
Speak slowly	SPS
Stammer	STAM
Content words	
Afraid	Afra
Anger	Anger
Angry	Angry
Annoy	Annoy
Anxiety	Anxiet
Argue	Argue
Boring	Bor
Bothering	Bother
Comfort	Cfor
Confidence	Cfid
Conflict	Cfl
Confused	Cfuse
Corrupted	Corrup
Cry	Cry
Dark	Dark
Dead	Dead
Death	Death
Decadence	Decad
Delight	Delight
Depression	Depres
Despair	Desp
Dispirited	Dispir
Distressed	Distre
Downcast	Downc
Dreary	Dreary
Escape	Esca
Exhausted	Exhau
Fear	Fear
Flee	Flee
Frenzied	Frenz
Friend	Frie
Frustrated	Frustr
Future	Future
Glad	Glad
Gloom	Gloom
Grief	Grief
Grieved	Grieved
Happy	Happ
Hate	Hate
Hope	Hope
Impulsion	Impuls
Indignant	Indign
Intimacy	Intima
Irritating	Irrita
Joyous	Joyo
Lacrimation	Lacri
Like	Like
Lonely	Lone
Loss	Loss
Me	Me
Mom	Mom
Mother	Mother
Nice	Nice
Numb	Numb
Oppressing sensation	OpSen
Optimism	Optimi
Pain	Pain
Parents	Pare
Pathetic	Pathe
Patient	Patient
Pessimism	Pessim
Pleasant	Pleasa
Quarrel	Quarr
Repression	Repres
Sad	Sad
Satisfaction	Satisf
Sensitive	Sensi
Shiver	Shiver
Slump	Slump
Somber	Somb

Sorriness	Sorrin
Sorrow	Sorrow
Stimulated	Stimul
Stress	Stres
Stubborn	Stubb
Suffocative	Suff
Suicide	Suic
Testiness	Testin
Troubled	Troubl
Uninteresting	Unint
Unpleasant	Unple
Weary	Weary
Whiny	Whiny

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40359-025-02527-0.

Supplementary Material 1.

Supplementary Material 2: Table 1 The matrix of co-occurrence probability coefficients indicating associations between nonverbal behaviors and content words based on the annotation frequency in subclinical group (in part). Table 2 The matrix of co-occurrence probability coefficients indicating associations between nonverbal behaviors and content words based on the annotation duration in subclinical group (in part). Table 3 The matrix of co-occurrence probability coefficients indicating associations between nonverbal behaviors and content words based on the annotation frequency in control group (in part). Table 4 The matrix of co-occurrence probability coefficients indicating associations between nonverbal behaviors and content words based on the annotation duration in control group (in part). Table 5 The matrix of co-occurrence probability coefficients indicating associations of nonverbal behaviors with vocal emotion and prosody based on the annotation frequency in subclinical group (in part). Table 6 The matrix of co-occurrence probability coefficients indicating associations of nonverbal behaviors with vocal emotion and prosody based on the annotation frequency in control group (in part). Table 7 The matrix of co-occurrence probability coefficients indicating associations of nonverbal behaviors with vocal emotion and prosody based on the annotation duration in subclinical group (in part). Table 8 The matrix of co-occurrence probability coefficients indicating associations of nonverbal behaviors with vocal emotion and prosody based on the annotation duration in control group (in part).

Acknowledgements
Not applicable.

Authors' contributions
LW and HZ developed the study concept and contributed to study design. LW implemented the experiment and collected data. LW, FW, and DL analyzed the data. LW and HZ wrote and revised the manuscript.

Funding
This study was supported by the National Social Science Fund of China (grant number 20BYY071).

Data availability
The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate
This study received ethical approval from the Human Research Ethics Committee of Jimei University (# JMU202405058). Our study complies with the Declaration of Helsinki. Informed consent was obtained from all subjects.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Teachers College, Jimei University, Jimei District, 185 Yinjiang Rd, Xiamen 361021, P.R. China. [2]School of Software, Jiangxi Normal University, Nanchang 330022, P.R. China. [3]School of Foreign Languages, Jimei University, Jimei District, 185 Yinjiang Rd, Xiamen 361021, P.R. China. [4]School of Computer Science & Technology, Beijing Institute of Technology, Haidian District, 5 Zhongguancun Nan Dajie, Beijing 100081, P.R. China.

Received: 18 September 2024 Accepted: 21 February 2025
Published online: 06 March 2025

References

1. Holler J, Levinson SC. Multimodal language processing in human communication. Trends Cogn Sci. 2019;23(8):639–52.
2. Serré H, Dohen M, Fuchs S, Gerber S, Rochet-Capellan A. Speech breathing: variable but individual over time and according to limb movements. Ann N Y Acad Sci. 2021;1505(1):142–55.
3. Yu K, Zhou Y, Liu B, Cai H, Wang R. Listeners' processing of linguistic and paralinguistic information in speakers' voice (in Chinese). Psychol Res. 2021;14(1):29–36.
4. Zhang Y, Frassinelli D, Tuomainen J, Skipper JI, Vigliocco G. More than words: Word predictability, prosody, gesture and mouth movements in natural language comprehension. Proc Biol Sci. 1955;2021(288):e20210500.
5. Beck AT. A 60-year evolution of cognitive theory and therapy. Perspect Psychol Sci. 2019;14(1):16–20.
6. Beck AT, Bredemeier K. A unified model of depression: Integrating clinical, cognitive, biological, and evolutionary perspectives. Clin Psychol Sci. 2016;4:596–619.
7. Sass K, Habel U, Kellermann T, Mathiak K, Gauggel S, Kircher T. The influence of positive and negative emotional associations on semantic processing in depression: a fMRI study. Hum Brain Mapp. 2014;35(2):471–82.
8. Chen Y, Li S, Guo T, Xie H, Xu F, Zhang D. The role of dorsolateral prefrontal cortex on voluntary forgetting of negative social feedback in depressed patients: a TMS study (in Chinese). Acta Psychol Sin. 2021;53(10):1094–104.
9. Zhou H, Dai B, Rossi S, Li J. Electrophysiological evidence for elimination of the positive bias in elderly adults with depressive symptoms. Front Psych. 2018;9:e62.
10. Semkovska M, Quinlivan L, O'Grady T, Johnson R, Collins A, O'Connor J, et al. Cognitive function following a major depressive episode: a systematic review and meta-analysis. Lancet Psychiatry. 2019;6(10):851–61.
11. Balsters MJH, Krahmer EJ, Swerts MGJ, Vingerhoets AJJM. Verbal and nonverbal correlates for depression: a review. Current Psychiatry Reviews. 2012;8:227–34.
12. Liu J, Shi M. What are the characteristics of user texts and behaviors in Chinese depression posts? Int J Environ Res Public Health. 2022;19(10):1–13.
13. Bernard JD, Baddeley JL, Rodriguez BF, Burke PA. Depression, language, and affect: an examination of the influence of baseline depression and affect induction on language. J Lang Soc Psychol. 2016;35(3):317–26.
14. Stirman SW, Pennebaker JW. Word use in the poetry of suicidal and nonsuicidal poets. Psychosom Med. 2010;63:517–22.
15. Cacheda F, Fernandez D, Novoa FJ, Carneiro V. Early detection of depression: social network analysis and random forest techniques. J Med Internet Res. 2019;21:e12554.
16. Huang G, Zhou X. The linguistic patterns of depressed patients (in Chinese). Adv Psychol Sci. 2021;29(5):838–48.
17. Smirnova D, Cumming P, Sloeva E, Kuvshinova N, Romanov D, Nosachev G. Language patterns discriminate mild depression from normal sadness and euthymic state. Front Psych. 2018;9:e105.
18. Xu, S, Yang, Z, Chakraborty, D, Victoria Chua, YH, Dauwels, J, Thalmann, D, et al. Automated Verbal and Non-verbal Speech Analysis of Interviews of

Individuals with Schizophrenia and Depression. Conference proceedings (IEEE Engineering in Medicine and Biology Society. Conf.) 2019:225–8. https://doi.org/10.1109/EMBC.2019.8857071.
19. Guo W, Yang H, Liu Z, Xu Y, Hu B. Deep neural networks for depression recognition based on 2D and 3D facial expressions under emotional stimulus tasks. Front Neurosci. 2021;15:e609760.
20. Michalak J, Troje NF, Fischer J, Vollmar P, Heidenreich T, Schulte D. Embodiment of sadness and depression-gait patterns associated with dysphoric mood. Psychosom Med. 2009;71:580–7.
21. Dehcheshmeh, TF, Majelan, AS, Maleki, B. Correlation between depression and posture (A systematic review). Curr Psychol. 2023; 1–11. https://doi.org/10.1007/s12144-023-04630-0.
22. Li W, Wang Q, Liu X, Yu Y. Simple action for depression detection: using kinect-recorded human kinematic skeletal data. BMC Psychiatry. 2021;21(1):1–11.
23. Williamson, J, Godoy, E, Cha, M, Schwarzentruber, A, Khorrami, P, Gwon, Y, et al. Detecting depression using vocal, facial and semantic communication cues. Paper presented at the AVEC '16: Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge. 2016:11–18. https://doi.org/10.1145/2988257.2988263.
24. Pan W, Flint J, Shenhav L, Liu T, Liu M, Hu B, Zhu T. Re-examining the robustness of voice features in predicting depression: compared with baseline of confounders. PLoS One. 2019;14(6):e218172.
25. Silva WJ, Lopes L, Galdino MKC, Almeida AA. Voice acoustic parameters as predictors of depression. J Voice. 2024;38(1):77–85.
26. Wang J, Zhang L, Liu T, Pan W, Hu B, Zhu T. Acoustic differences between healthy and depressed people: a cross-situation study. BMC Psychiatry. 2019;19(1):e300.
27. Garcia-Toro M, Montes JM, Talavera JA. Functional cerebral asymmetry in affective disorders: new facts contributed by transcranial magnetic stimulation. J Affect Disord. 2001;66(2–3):103–9.
28. Rodríguez MR, Nuevo R, Chatterji S, Ayuso-Mateos JL. Definitions and factors associated with subthreshold depressive conditions: a systematic review. BMC Psychiatry. 2012;12:181.
29. Sun C, Yan C, Lv Q, Wang Y, Xiao W, Wang Y, Yi Z, Wang J. Emotion context insensitivity is generalized in individuals with major depressive disorder but not in those with subclinical depression. J Affect Disord. 2022;313:204–13.
30. Wang Z, Yuan C, Huang J, Li Z, Chen J, Zhang H, et al. Reliability and validity of the Chinese version of beck depression inventory-II among depression patients (in Chinese). Chin Ment Health J. 2011;25(6):476–80.
31. Spitzer RL, Kroenke K, Williams JBW. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. JAMA. 1999;282(18):1737–44.
32. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606–13.
33. Chen M, Sheng L, Qu S. Diagnostic test of screening depressive disorders in general hospital with the patient health questionnaire (in Chinese). Chin Ment Health J. 2015;29(4):241–5.
34. Lausberg H, Sloetjes H. Coding gestural behavior with the NEUROGES-ELAN system. Behav Res Methods Instrum Comput. 2009;41(3):841–9.
35. Assenov Y, Ramírez F, Schelhorn SE, Lengauer T, Albrecht M. Computing topological parameters of biological networks. Bioinformatics. 2008;24(2):282–4.
36. Wang L, Zhang H, Lin D. Multimodal analysis of vocal emotion and prosody in subclinical depressed individuals. Curr Psychol. 2025. https://doi.org/10.1007/s12144-025-07313-0.
37. Wang L, Sang B. Embodied cognition of conceptes (in Chinese). J Nantong Univ (Soc Sc Edition). 2014;30(4):100–6.
38. Wang L, Ke J, Zhang H. A functional near-infrared spectroscopy examination of the neural correlates of mental rotation for individuals with different depressive tendencies. Front Hum Neurosci. 2022;16: e760738.
39. Wu G. On the origin of thinking and language (in Chinese). Soc Sci China. 1981;3:25–40.
40. Gu Y. On the consistency principle among word-mind-emotion-appearance and live language (in Chinese). Contemporary Rhetoric. 2013;6:1–19.

Publisher's Note