Output Analysis in Voice Interaction in AI Environment

Fanyu Jin
Yancheng Institute of Technology, Yancheng224051, Jiangsu Province, China
E-mail: tomjin2001@126.com

The future foreign language teaching will inevitably be combined with AI technology, and it is likely that the traditional foreign language teaching method of one teacher instructing a number of students will gradually be completely replaced by a new foreign language learning mode of each student's foreign language learning and most importantly acquisition being realized by interacting with AI interface customized for each student. The reason is rather simple. In spite of numerous repatching teaching method explorations, the traditional classroom-based foreign language teaching has been unable to solve the congenital problems, such as inadequacy of language input and interaction, insufficiency of real life sensory stimulation and violation of natural language acquisition sequence for the absence of language environment. The AI interface, in contrast, with the infinite, accurate and real language supply and human-computer interaction, and also with constant adjustment of ZPD (Zone of Proximal Development) according to each student's language development level, precisely sets up appropriate scaffolding for every language learner, thus revolutionarily creating a language environment close to or even beyond the real one and returning language learning to natural acquisition process. The first step to achieve this goal is to realize human-computer voice interaction. The realization of voice interaction needs many technical supports, among which voice interaction output analysis is an urgent part. By importing AI voice interactive output analysis algorithm, constructing output analysis model, and establishing the operation platform of the analysis model, the paper relies on the determination of the voice interactive output influence function, and takes the cultural elements of English language as an example to analyze the output.

Povzetek: V članku je objavljana analiza govorne komunikacije z upoštevanjem mehanizmov angleškega kulturnega okolja.

1 Introduction

Interaction based on mechanical one-way input have been giving way to the two-way voice interaction in numerous ways to satisfy human-computer communication needs especially in language learning. For instance, on the basis of Google search function, Google Now records the keywords searched by users, and provides users with relevant voice services through intelligent reading. This allows the machine to upgrade from "passive" answering user's questions to "active" alerting users to their needs, that is, the way the machine interacts with human beings in a service-oriented manner. Whether it's Apple AI or Google Now, it gives machines the ability to act on the basis of "independent thinking", thus opening a new era of language learning with two-way human-computer interaction. Conventional voice interactive output analysis method uses dynamic voice capture technology to realize voice interactive output analysis, which can greatly improve the efficiency of voice output. However, when applied to voice interactive output analysis in Artificial Intelligence (AI) environment, due to the high degree of strangeness in the field and the limited response of the operating environment, the problem of low accuracy of output analysis appears. This greatly reduces the frequency of AI environment use, resulting in incomplete output analysis of voice interaction in AI environment. In order to ensure the validity of the interactive output analysis method and simulate the language interaction environment in AI environment, two different methods of interactive output analysis are used to analyze the output accuracy simulation experiments. It is found that the method in this paper has higher analysis output accuracy [1].

2 Construction of AI voice interactive output analysis model

At present, the internationally recognized and accepted communicative teaching method aims to cultivate students' ability to express themselves in the target language. Interactive teaching method is a kind of communicative language teaching. Interactive teaching theory holds that language is the system of expressing ideological system, and the main function of language is interaction and communication. Interaction mainly refers to the interaction between teachers and students and also among students themselves in the classroom. Classroom activities should include real communication and enable students to perform meaningful tasks. Students'

communication includes sharing information and negotiating meanings with others.

The construction of AI voice interactive output analysis model mainly includes two parts: building the running platform of AI voice interactive output analysis model and importing AI voice interactive output analysis algorithm.

2.1 Establishment of AI voice interactive output analysis model platform

The model of *Multi-dimensional Input-interactive Output* is based on the premise of "language learning is technical training, not pure knowledge learning", and guided by Robert W. Blair's *Low Shielding Effect* in 1987 and Swain's *Output Hypothesis* in 1995. In the process of building this model, we will focus on the real language input mode and the learner's output mode which play decisive roles in the latter language output.

The running platform of AI voice interactive output analysis model is the basic platform to ensure the reasonable and safe operation of AI voice interactive output analysis model. The platform consists of four parts: data layer, operation layer, physical layer and display layer [2].

Data layer is the logical level for acquiring voice interactive information, which provides data support and data guarantee for operation layer. Under the premise of authorization, the security environment is guaranteed, the voice interaction status and voice interaction information are acquired, and the data supply is completed.

Operational layer is a logical level based on AI voice interaction output analysis algorithm, which provides direct and indirect evidence for output analysis in AI voice interaction.

The physical layer includes all external devices supporting the platform, such as processors, hosts, displays, network connections, etc. It provides a hardware platform for the platform and output analysis, in AI environment.

The display layer is the logical level through which the results are displayed by the display devices in the physical layer, so that the staff can read the results directly and realize the output analysis in the voice interaction in AI environment [3].

2.2 An interactive output analysis algorithm for importing AI speech

For the studying the output analysis of voice interaction in AI environment, the algorithm of AI voice interaction output analysis' main function is to analyze the speech interaction output of artificial intelligence. In the environment of voice interaction, taking the cultural elements of English language as an example, AI voice interactive output analysis algorithm can not only drive mathematical operations, but also complete the transforming voice into written language. In the process of data calculation, the conversion process from voice to language and the process of driving mathematical operation are analyzed in detail.

In the process of voice-to-language transformation, the parameters are generated and the results are output mainly through voice acquisition module and voice recognition module. With different results of the output, the command is written in the form of MySQL statement. The key program of the command is shown in Figure 1.

```
· int *q,*#include<stdio.h>
· Llist(size)S.base=(char*) malloc(1005* sizeoth)) return 0;
· Int *p,*q;(S.top==charfor(i=0;i<strlen(str);i++) f(char));if((i<1)||(i>L.leng
· New base=(int*)realloc(L.elem,(L.listsize+50)*sizeof(int))  ;elete(Sq List &L,int I,int
  &e)return  *S.top++=e;S.size=1005ile(
· If(!new base) ex it·  Lelem=neif(str[i]<='z(char));
· q=&L.el=fopen("test","w"))==NULL)char S.",&n)! get;(-2);wh
· *q=e;ifrintf("error: cannot open file!\n");=r)&&str[i]>='a')5*sizeof
· L.length++;return 1n--) if((i<1)||(i>L.length)) return 0;base) ("%d
· end
```

Figure 1: Key program of AI voice interactive output analysis algorithm.

3 Output analysis in implementing voice interaction in AI environment

Based on the construction of AI voice interactive output analysis model, the influence function of voice interactive output is determined. Taking the cultural elements of English language as an example, the output analysis is carried out to realize the output analysis of voice interaction in AI environment.

To determine the influence function of voice interactive output, we need to determine the state equation of voice interaction and calculate the interaction coefficient of voice, so as to realize the determination of the influence function of voice interactive output.

The speech interaction state is real-time interaction, and is difficult to quantify, thus it is necessary to construct the speech interaction state equation. Different types of language interaction have different operational systems of equations and methods of constructing equations. This paper takes the voice interaction between English users as an example to construct the state equation of voice interaction.

Based on the relevant information acquisition module of AI voice interactive output analysis model operation platform, the AI voice interactive output analysis algorithm is imported for statistical analysis. Assuming that the amount of interactive information is M and the interaction coefficient q is 1.0, there is a state equation of inter-cell voice interaction, as is shown in formula (1)[4].

$$C = -\frac{q_x}{\left(\dfrac{\partial T}{\partial M}\right)} \quad (1)$$

In the formula, C represents the state equation of inter-cell voice interaction, and q_x represents the type of language, and T represents the time of voice output. When the interaction mode is more complex, the interaction coefficient q can be expressed by formula (2):

$$q = Q / (C \cdot VT) \quad (2)$$

In the formula, Q represents the ideal coefficient, C represents the state equation of inter-cell voice interaction, and $\triangle T$ represents the type of interaction mode.

Through the determination of the speech interaction state equation, the speech interaction coefficient is calculated relying on the variable relationship of the speech interaction state equation. The calculation of speech interaction coefficient is based on the interaction equation. The calculation process is as follows[5]:

$$\rho = q \lim_{\Delta V \to 0} \frac{F^2}{VD} \quad (3)$$

In the formula, D represents the interactive state, V represents the interactive output, and F represents the scope of language interaction[6]. When the speech interaction coefficient does not satisfy the interactive equation, the statistical calculation of the data is carried out, as is shown in formula (4).

$$\rho = C \lim_{\Delta V \to 0} \frac{F^2}{D} \quad (4)$$

The condition that the speech interaction coefficient does not satisfy the interaction equation is as follows: Speech characters are less than 4 characters, there are uncommon words in the speech which cannot be recognized, and the speech discrimination is not high, making it difficult to recognize[7].

Based on the calculation of voice interaction coefficient, the influence function of voice interaction output is determined, which can be expressed by formula (5):

$$E = -\rho k \left(\frac{\partial g}{\partial f} \right)^2 \quad (5)$$

In the formula, g is an interactive way, such as English. f is running, and k represents the reliability of interactive running program[8].

The determination of the influence function of the speech interactive output analysis algorithm is realized by the formulas (1)~(5).

4 Example analysis

In order to ensure the output analysis in the speech interaction in AI environment proposed in this paper, the cultural elements in the English language are used as examples, and the English audio interaction in different AI environments is used as the test object to analyze the output precision simulation experiment[9]. Different cultural forms of English sound interaction in AI environment, speech types, dialects, etc. are simulated [10]. The simulation experiment was carried out by using the conventional interactive output analysis method as the experimental comparison object[11].

4.1 Test data preparation

In order to ensure the accuracy of the analysis output simulation test, the experimental data preparation is first carried out, and the SDH-214 simulation system is selected for the data operation platform [12]. On the two sets of identical computers, two different interactive output analysis methods were used to analyze the output precision simulation [13]. The experiment mainly consisted of English speech interaction, with different English customs and cultural characteristics for analysis[14].

4.2 Test results analysis

During the experiment, two different methods of interactive output analysis were used to analyze the change of output accuracy in the simulation environment. The simulation curves of analytical output accuracy are obtained, as is shown in Figure 2.
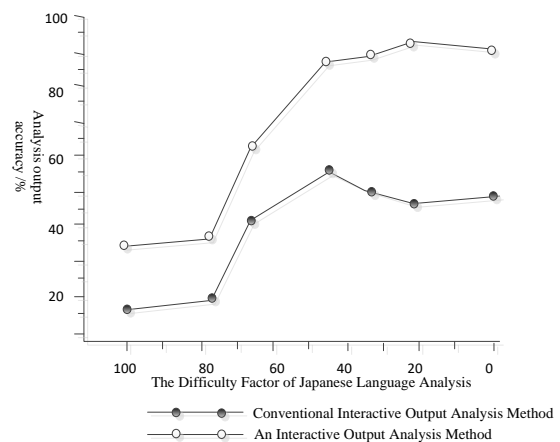

Figure 2: The output accuracy simulation curve.

According to the analysis of the test curve results, the output accuracy of the proposed interactive output analysis method is 73.45%, and that of the conventional interactive output analysis method is 39.24%. Compared with the conventional interactive output analysis method, the output accuracy of the proposed interactive output analysis method is 34.21%, which is suitable for the output analysis of voice interaction in AI environment [15].

5 Conclusion

This paper presents the output analysis of voice interaction in AI environment with examples of cultural elements in language. The research conducted on the base of the construction of AI voice interactive output analysis model and the determination of relevant parameters. The experimental data show that the proposed interactive output analysis method has high analysis output accuracy. This study provides a new idea for the interactive output analysis method, as well as a theoretical basis for the interactive output analysis method, and lays a foundation for the further study of voice interaction analysis. However, there are still some deficiencies in the practical application of this paper. The author hopes to further improve the output accuracy of voice interaction in the future research.

6 References

[1] KimHC . Weaknesses of Voice Interaction[C]// International Conference on Networked Computing & Advanced Information Management. *IEEE*, 2008.

[2] KostovV , FukudaS . Emotion in user interface, voice interaction system[C]// IEEE International Conference on Systems. *IEEE*, 2000.

[3] Shriver S ,TothA , ZhuX , et al. A Unified Design for Human-Machine Voice Interaction[C]// Chi 01 Extended Abstracts on Human Factors in Computing Systems. *ACM*, 2001.

[4] Osawa H , Orszulak J , Godfrey K M , et al. Improving voice interaction for older people using an attachable gesture robot[C]// Ro-man. *IEEE*, 2010.

[5] MassieT ,WijesekeraD . TVIS: Tactical Voice Interaction Services for dismounted urban operations[C]// Milcom IEEE Military Communications Conference. *IEEE*, 2013.

[6] LeeA ,OuraK , TokudaK . Mmdagent—A fully open-source toolkit for voice interaction systems[C]// IEEE International Conference on Acoustics. *IEEE*, 2013.

[7] DrigasA ,ArgyriK , VrettarosJ . Decade Review (1999-2009): Artificial Intelligence Techniques in Student Modeling.[C]// Best Practices for the Knowledge Society Knowledge, Learning, Development & Technology for All, Second World Summit on the Knowledge Society, Wsks, Chania, Crete, Greece, September. *DBLP*, 2009.

[8] GorostizaJF ,SalichsMA . Teaching sequences to a social robot by voice interaction[C]// IEEE International Symposium on Robot & Human Interactive Communication. *IEEE*, 2009.

[9] WolfgangGarn,JamesAitken. Agile factorial production for a single manufacturing line with multiple products[J]. *European Journal of Operational Research*,2015(3).

[10] JamesAitken,CecilBozarth,WolfgangGarn. To eliminate or absorb supply chain complexity: a conceptual model and case study[J]. *Supply Chain Management: An International Journal*,2016(6).

[11] AfshinSamani,DivyaSrinivasan,SvendErikMathiassen,PascalMadeleine. Nonlinear metrics assessing motor variability in a standardized pipetting task: Between- and within-subject variance components[J]. *Journal of Electromyography and Kinesiology*,2015(3).

[12] Matsushita Y , Uchiya T , Nishimuray R , et al. Crowdsourcing Environment to Create Voice Interaction Scenario of Spoken Dialogue System[C]// International Conference on Network-based Information Systems. *IEEE*, 2015.

[13] Zhou F Y , Li J H , Tian G H , et al. Research and Implementation of Embedded Voice Interaction System Based on ARM in Intelligent Space[J]. *Advanced Materials Research*, 2012, 433-440:5620-5627.

[14] MitreaD ,Mitrea M . Voice interaction in an integrated office and telecommunications environment[J]. *J.am.chem.soc*, 2015, 111(8):1138-1157(20).

[15] Coppola P , Mea V D , Gaspero L D , et al. AI Techniques in a Context-Aware Ubiquitous Environment[M]// *Pervasive Computing*,Springer London, 2009.