

Exploring a Model for Training Higher Vocational College Students' Direct Processing in English Listening: A Working Memory Perspective

Liu Yanping*

Guangdong Institute of Arts and Sciences, Lianjiang, Guangdong, China

**Corresponding Author*

Abstract: The "translation during listening", which has caused listening comprehension failure and low English proficiency among many higher vocational college students, has been rarely explored. This study, from the working memory perspective, analyzes the cause of this "translation" phenomenon and proposes a model with five strategies to help students remove the "translation" link and build direct sound-meaning mapping in English listening.

Keywords: Higher Vocational College English; Working Memory; Direct Processing

1. Introduction

Listening plays a foundational role in English learning. In listening class, the author has observed a phenomenon among higher vocational college students, especially those with weak English foundations: listen → translate into Chinese → understand. This "translation" habit not only causes students' listening comprehension failure, but also hinders listening teaching efficiency, ultimately leading to failure in improving their language ability. Listening is a process related to cognitive processing, not just language knowledge. This so-called "translation during listening" phenomenon is closely related to how students allocate limited cognitive resources in various stages during listening and how they deal with the mapping relationship between phonological and semantic information.

2. Literature Review

According to the China National Knowledge Infrastructure, many studies on English listening teaching in higher vocational colleges has been conducted around such themes as current situation surveys with countermeasures, technologies application and teaching models. For example, Sun Yijia (2025) explored the

application of digital tools in English listening teaching and Zhang Yi (2025) explored the obstacles and countermeasures in English listening for students in higher vocational colleges. But very few of these studies delve into the "translation during listening" phenomenon. Although a few studies has applied the working memory theory, they either focus on reading instruction (Wu Yuan, 2011) or grammar instruction (Zhao Junxia et al., 2020). Very few related to listening and working memory can be found, as seen in Zheng Jiasheng's (2012) exploration of the facilitating effect of dictation exercises on English-major students' listening ability from the working memory perspective. To sum up, this "translation during listening" phenomenon has not been thoroughly delved into, especially from the working memory theory perspective to explain why it leads to students' listening comprehension failure and how to tackle it.

3. Working Memory and Listening Comprehension

Working memory, an important cognitive concept proposed by Baddeley and Hitch in 1974, refers to the system for temporarily storing and processing information when individuals perform complex cognitive tasks. It differs from short-term memory as it emphasizes active processing rather than passive retention. Baddeley (2000) refined the working memory model with three subsystems. The central executive oversees attention deployment, cognitive coordination, and cross-subsystem information flow, essentially a resource manager that prioritizes competing demands. The phonological loop, the foundational channel for connecting sound and meaning, manages the temporary storage and rehearsal of phonological information. Phonological storage retains phonological traces that rapidly decay while phonological rehearsal refreshes this stored

information through subvocal repetition to prevent information loss. The visuo-spatial sketchpad serves as a parallel function for visual and spatial content, managing images, layouts, and spatial arrangements. According to this model, working memory capacity is limited. When information exceeds the cognitive load of working memory, the cognitive processing would fail and information would be under-processed or lost. This is the key to explaining students' listening comprehension failure as there is a positive correlation between working memory capacity and L2 listening performance.

The "listen → translate into Chinese → understand" process can be explained in terms of working memory. When students adopt a "translation strategy" in listening comprehension, the cognitive processing flows like: listen to English sounds → activate English vocabulary → retrieve Chinese equivalents → assemble a Chinese sentence → understand the Chinese meaning. Compared with the direct processing (listen to English sounds → activate English vocabulary → understand meaning), but this strategy includes two extra links: retrieve Chinese equivalents and assemble a Chinese sentence.

The problem is that these two extra links consume precious cognitive resources. As the working memory capacity is limited, when cognitive resources are heavily occupied by the translation link, insufficient resources remain for phonological retention and semantic integration. Thus, as incoming phonological information keeps arriving, the earlier input remains partially or improperly processed. So when the capacity exceeds, information losses, causing comprehension failure, which is what happen in students' mind when they pause midway in listening, seemingly lost in thought.

So, to develop students' direct processing in listening is the solution. That's, to help students' get over the "translation" link and establish direct mapping between phonological input and semantic meaning to achieve direct sound-meaning response. If they can remove this intermediary link between listening to English and understanding English, the cognitive load on their working memory would be reduced. So, more cognitive resources can be used to retain and integrate information, thus improving their listening comprehension.

4. Construction of the Direct Processing Training Model

4.1 Model Overview and Aims

This model, based on the theoretical framework above, aims to help higher vocational college students with weak English foundations to build a direct sound-meaning connection to overcome the "translation during listening" habit, relieving their working memory load during listening. The design logic is: starting from the limited capacity of working memory, then moving to an analysis of the "translation" strategy's cognitive consumption, then to training activities that foster direct mapping, and finally to consistent practice that consolidates automatic processing. It is expected to achieve compressed cognitive load and improved listening efficiency through restructuring how students process incoming speech.

This model has three progressive aims. The immediate aim is to help students reduce Chinese intervention and establish a direct mapping between English sound and meaning. The deeper aim is to compress students' cognitive load on working memory during listening, enhancing their immediate listening response and information retention. The ultimate aim is to develop students' habit of thinking directly in English, achieving a shift from "learning English" to "thinking in English." The immediate aim is the starting point, while the deeper aim is the key, and the ultimate one is the orientation.

4.2 Teaching Principles

Meaning Priority. The purpose of listening training is to help students grasp meaning, instead of word-by-word identification. In teaching, students should be allowed to grasp the main idea or key information of the listening material instead of focusing on 100% word identification. Because in listening, if students pay too much attention to the meaning of every word, their working memory would be overly consumed by the low-level task (like phonological identification), leading to failure in high-level task (meaning integration).

Speed Priority. In listening class, students often pause to translate mentally. This breaks the listening and information flow. So, time pressure in training is introduced. For instance, requiring a response within three seconds after a listening part ends. This pushes students out of the "listen

→ translate → understand" habit and forces them to attempt the more direct "listen → understand" route, so their cognitive processing would be rushed into a faster and more automatic mode.

Gradual Withdrawal of Support. Early in the training, don't throw students into the deep end. Use temporary supports like videos, pictures, or keyword prompts to ease them in by reducing their processing burden. When students gain confidence and accuracy, pull them back one by one. This is to gradually push students' brain into relying on phonological information alone, without the help mentioned above.

Strategic Anticipation. Listening is inherently uncertain as speech varies in speed, accent, etc. To reduce uncertainties, equip students with practical moves: context based prediction, keywords catching, and selective attention to what matters most. Each strategy essentially reduces the possibilities their brain must entertain. Less uncertainties mean fewer cognitive demands on the central executive, which in turn frees up cognitive resources for meaning integration. In short, don't just train students' ears; train their minds to listen more intelligently.

4.3 Specific Strategies

Listening Prediction Training. Before listen comprehension, offer students cues like titles, pictures, or background information to predict what they will listen to. Use group discussion: students share their predictions, then listen and check if they were right. Prediction activates what cognitive psychologists call content schemata (students' previous relevant knowledge). When students hear something that matches their expectation, their working memory just needs to confirm it instead of processing it from scratch, which saves them processing capacity. More importantly, prediction turns listening from a passive receiving activity into an active checking activity, and that shift alone increases both students' attention and engagement.

Immediate Response Tasks. Design simple tasks that must be completed within 3-5 seconds after listening, such as True/False judgments, choosing corresponding pictures, performing physical actions, or completing a half-sentence fill-in-the-blank. The tasks must be simple, not allowing students time to translate as time pressure is very effective to interrupt the

translation chain. As students must react within an extremely short time, they lack time to complete the full chain of "listening → translating → understanding → answering" and can only resort to "listening → understanding → answering." Repeated training will reinforce this shortcut, shifting it from a conscious effort to automatic processing. Once the direct sound-meaning mapping reaches an automatic level, the burden on students' working memory will significantly decrease.

Shadowing Practice. Play the listening materials (starting at a slow speed), and students repeat (shadowing) what they hear with a delay of about half a second to one second, imitating the original pronunciation, intonation, pauses, and rhythm as closely as possible. Students begin with transcripts, shadowing the speaker's words as they read along. Over time, take the transcripts away, leaving only the audio. This forces students to hold the speech in their phonological loop long enough to repeat it back. Why does it help? First, shadowing gives the phonological loop a rigorous workout: the more the training, the longer sound traces persist, giving higher-level semantic processing more raw material to work with. Second, shadowing pushes students' brain to chunk speech into meaningful units rather than processing isolated words, which in turn reduces the moment-by-moment load on working memory.

Keyword Filtering. Train students to selectively pick keywords (nouns, verbs, adjectives, etc.) and neglect filler words or redundant information. In the early stage, set the goal of telling "the number of key words" (for example, "After listening, tell me which 5 key words you've heard"), and later transit to "retelling the main idea with the key words". The secret lies in the constraint of working memory. Information input in listening is continuous, but working memory can't process all the information simultaneously. Thus, keyword capture training can help students recognize the functions of selective attention. That's, students focus their limited cognitive resources on the language components with the largest amount of information and the most crucial factor for meaning construction. This means a preprocessing of the information input and a reduction of the total information entering the working memory, so students can allocate their limited cognitive resources to the core content.

Phased Withdrawal of Visual Scaffolding.

Approach this with a step-up routine. For a given listening material, move through three phases: first, students watch a video with subtitles; second, they listen to the audio with keyword prompts; third, they listen to the audio with no support at all. Students repeat each phase as needed, and only move them up when they are ready. The secret lies in the scaffolding theory. Provide support early, then gradually pull it back. Visual aids like subtitles and keywords ease students' initial burden: the visual channel supplements the auditory, reducing the pressure on phonological identification. But if these aids are kept too long, students never learn to process sound independently. The withdrawal process forces the auditory channel to take over, progressively building its own capacity for sound-meaning mapping until it can stand alone.

4.4 Teaching Process Design

The five strategies above can be integrated into a single classroom session. Here is how a typical 20- to 40-minute listening lesson might flow.

Pre-listening Stage (5-8 min): Prediction activity: Present topic keywords/pictures, students discuss and predict content. This stage is to activate students' schemata and reduce uncertainty.

While-listening Stage: first listening (5-8 min): the teacher plays material and students complete immediate response tasks (True/False, picture selection, etc.), so as to force students' rapid response, break the translation chain; second listening (5-8 min): the teacher plays material and students capture and record keywords to train students' selective attention to compress information volume; third listening (5-8 min): the teacher plays material (slower speed) and students perform shadowing to strengthen students' phonological loop and help establish sound-meaning mapping.

Post-listening Stage (5-8 min): Group discussion to verify predictions; the teacher explains key and difficult points with the aim of consolidating students' understanding and confirming sound-meaning correspondences.

This model targets higher vocational college students with weak English foundations who habitually translate during listening. Implementation should begin with very simple tasks, gradually increasing difficulty to avoid triggering students' anxiety. Consistent practice is essential. Immediate feedback is required to reinforce students' correct sound-meaning mappings. Differentiated tasks should be

designed to accommodate varying proficiency levels.

5. Conclusion

This is a tentative study exploring the "translation during listening" phenomenon among many higher vocational college students from the working memory perspective, and discovers that this "translation" habit consumes students' limited cognitive resources, causing reduced cognitive resources for phonological retention and semantic integration and ultimately listening comprehension failure. To address this, a model is proposed to train students' direct processing, including five specific strategies (prediction training, immediate response tasks, shadowing practice, keyword filtering, and phased withdrawal of visual scaffolding). This model is still at the exploratory stage and future research may verify its effectiveness through teaching experiments and explore its differentiated applications at different language proficiency levels.

References

- [1] Baddeley, Alan. D., & Hitch, Graham. J. Working Memory. In Gordon H. Bower (Ed.), *Recent Advances in Learning and Motivation* (Vol. 8, pp. 47-89), 1974. New York: Academic Press.
- [2] Baddeley, Alan. D. The Episodic Buffer: A New Component of Working Memory?[J]. *Trends in Cognitive Sciences*, 2000, 4(11): 417-423.
- [3] Sun Yijia. Research on the Application of Digital Tools in Higher Vocational English Listening Teaching [J] *Overseas English*, 2025,(19): 213-215.
- [4] Wu Yuan. An Exploration of EFL Reading Teaching Method Based on Working Memory[J]. *Journal of Tianjin Vocational Institute*, 2011, 20(5): 70-72.
- [5] Zhao Junxia, Wang Jue, & Li Yanmin. Classroom Teaching Design for Higher Vocational College English Based Upon Cognitive Load Theory--Taking the Subjunctive Mood as an Example[J]. *Journal of Guangxi Vocational and Technical College*, 2020, 13(5): 86-91.
- [6] Zhang Yi. Analysis of English Hearing Impairment and Countermeasures among Students in Private Higher Vocational Colleges[J]. *Education and Teaching Forum*, 2025, (1): 185-188.

- [7] Zheng Jiasheng. A Study on the Effectiveness of Dictation Practice in Improving the Listening Ability of English

Major Students in Higher Vocational Colleges[J]. Journal of Changchun Education Institute, 2012, 28(11): 126+128.