
# Journal of Technology and Chinese Language Teaching

**Volume 10 Number 1, June 2019**
二〇一九年六月　第十卷第一期



**ISSN: 1949-260X**
**http://www.tclt.us/journal**

# 科技与中文教学

# Journal of Technology and Chinese Language Teaching

*A peer-reviewed online publication with in-print supplement*
*ISSN: 1949-260X http://www.tclt.us/journal*

**Volume 10 Number 1, June 2019**


## Articles



## Columns

曾妙芬 (Tseng, Miao-fen), *弗吉尼亚大学 (University of Virginia)*
高燕 (Gao, Yan), *亨利科县公立学校 (Henrico County Public Schools)*
蔡罗一 (Cai, Luoyi), *北卡罗莱纳大学教堂山分校, (University of North Carolina, Chapel Hill)*

科技与中文教学

# Journal of Technology and Chinese Language Teaching

*A peer-reviewed online publication with in-print supplement*
*ISSN: 1949-260X  http://www.tclt.us/journal*

## Sponsors

Department of Chinese Language and Literature, University of Macau
Department of World Languages, Literatures, and Cultures, Middle Tennessee State University
China Social Sciences Press (中国社会科学出版社)

## Editorial board



## Editorial staff

## Contacts

URL: http://www.tclt.us/journal
Email: editor@tclt.us

# 中文近义词辨析实验——机器学习程序与二语学习者的对比
# (An Experimental Study on Discriminating Chinese Near Synonyms: Contrasts between Machine Learning Systems and Second Language Learners)

詹卫东
(Zhan, Weidong)
北京大学
(Peking University)
zwd@pku.edu.cn

曹晓玉
(Cao, Xiaoyu)
北京大学
(Peking University)
xiaoyu.cao@pku.edu.cn

崔巍
(Cui, Wei)
北京大学
(Peking University)
cuiw@pku.edu.cn

常宝宝
(Chang, Baobao)
北京大学
(Peking University)
chbb@pku.edu.cn

**摘要：** 本文将机器学习技术引入中文近义词辨析任务，与二语学习者在近义词辨析任务上展开了初步的实验对比研究。在近义词集的选取、测试题制作方面，本文遵循平衡与周全原则。测试结果显示：机器在中文近义词辨析任务上的表现与二语学习者有明显可比性，机器测试成绩与中级水平的汉语学习者测试成绩呈正相关。除近义词本身难度有别外，题型差异对测试成绩有显著影响。词语意义特征对机器近义词辨析的影响并不低于形式特征的影响，在辨析时机器和二语者对句法形式特征的把握比对搭配区别特征的把握更有效。

**Abstract:** In this article, machine learning technology is introduced into fill-in-the-blank (FITB) tasks involving the discrimination of Chinese near synonyms. A preliminary experimental study was carried out on said tasks between machines and L2 learners of Chinese. This study adheres to principles of balance and comprehensiveness in selecting synonyms and making test sets for the experimental research. The test results show that the performance of machines in discriminating Chinese near synonyms in FITB tasks is significantly comparable to that of human L2 learners. The score of the machine was also positively correlated with that of intermediate-level Chinese learners. In addition to the sets of near synonyms varying in difficulty, the difference of test question types also has a significant impact on test scores. The influence of lexical meaning features on the discrimination of near synonyms is no less than that of its formal features. Meanwhile, it is more effective for machines and L2 learners to exploit syntactic formal features rather than distinguishing collocation features in FITB tasks.

**关键词：**近义词、易混淆词、机器学习、Bi-LSTM 模型、近义词辨析实验

**Keywords:** Near synonyms, confusable words, machine learning, Bi-LSTM, FITB tasks

## 1. 引言

作为语言教学的重要组成部分，词汇教学，特别是其中的近义词辨析，一直是受到研究者、教师和学习者普遍关注的一个领域，无论在理论研究层面还是在实践运用方面，都有诸多挑战（Nation & Newton, 1997、洪炜，2012）。随着基于语料库技术的实证型研究在对外汉语教学研究领域的普及，以及二语教学理论和实践的不断丰富和发展，面向对外汉语教学的近义词研究和实践近年来取得了非常多的成果，比如赵新、刘若云（2005）对编写面向二语学习者的近义词词典的思考；张博（2007）、高再兰（2016）对面向母语者的近义词辨析和面向二语学习者的近义词辨析关系的深入讨论；洪炜（2017）有关课堂显性辨析对二语学习者近义词辨析的积极影响的实验研究。Hong（2014）、陆方喆（2016）基于语料库开展的近义词辨析研究，等等。可以说，当前学界对近义词辨析的探讨既有宏观视角的开拓，也有微观技术手段的不断更新，呈现出多维度、多层面、多种技术手段百花齐放的蓬勃发展态势。

与此同时，近年来自然语言处理（NLP）在以深度学习（deep learning）为代表的新一代基于人工神经网络（Artificial Neural Network）的机器学习技术推动下，也有不少研究开始涉足机器近义词辨析和同义词可替换性的探索。这方面主要包括两类代表性的任务：一是选词填空任务（Fill-in-the-blank，简称 FITB），由 Edmonds（1997）提出。二是在国际语义评测会议[1]（SemEval 2007）上提出的词汇替换任务（Lexical Substitution Task，简称 LST），可参见 McCarthy & Navigli (2007、2009)、Zhao 等（2007）、陈士婷等（2012）。其中前一种任务的研究工作更多一些。NLP 研究人员尝试了不同的机器学习模型来完成 FITB 任务。Wang & Hirst (2010) 基于隐性语义分析法（Latent Semantic Analysis，LSA）和支持向量机（Support Vector Model，SVM）分类器，在 7 组英语近义词上进行选词填空任务实验，达到了 74.5%的准确率。Yu 等（2011）将上述 7 组英语近义词翻译为 7 组中文近义词进行了中文近义词辨析实验，采用点式互信息（PMI）和 5 元模型（5GRAM）在选自台湾中研院 Sinica 语料库和 Chinese News 语料库的测试集上进行辨析实验，准确率为 68.07%。Yu & Chien（2013）在 LSA 模型基础上增加了独立成分分析法（Independent Component Analysis）来寻取影响近义词辨析的隐性因素，进而基于 SVM 分类器完成近义词辨

[1] 国际语义评测会议（SemEval）已有 20 年历史，致力于组织与计算机语义分析相关任务的评测和学术交流。其前身是 Senseval（词义消歧评测）会议。可参见 https://en.wikipedia.org/wiki/SemEval。

析任务，在中、英文语料上的准确率分别达到 82.57%和 76.78%。特别值得注意的是，Huang 等（2017）首次将机器近义词辨析任务跟二语学习结合起来。该文对比了高斯混合模型（GMM）和双向长短时记忆神经网络（Bi-LSTM）模型，在英语 24 组同义词辨析实验任务上，两种模型的准确率分别达到了 78.16%和 83.59%，并进一步针对二语学习者开展了计算机辅助近义词学习实验，发现尽管基于 Bi-LSTM 模型的机器评测效果好于 GMM 模型，但是 Bi-LSTM 是对整个句子的全局区别把握有优势，GMM 则是对近义词相关的特定区别特征把握有优势，因此，基于 GMM 模型来辅助二语者的近义词学习，效果反而优于 Bi-LSTM 模型。

已有的研究显示，基于机器学习技术的近义词计算机自动辨析，有可能对二语学习提供帮助。这方面的工作才刚刚开始，是一个值得探索的研究方向，特别是在中文近义词的辨析方面，还缺乏机器学习实验与二语学习者实验的对比考察。本文从这个角度出发，希望将机器学习技术用于中文近义词辨析任务中，并在大致相当的环境下，考察机器、母语者、二语学习者对同一套测试题的表现是否有可比性，通过分析初步的实验结果，为今后进一步的深入研究寻找方向。下面第 2 节简要介绍本文实验所用的机器学习模型；第 3 节说明实验的设计；第 4 节是实验结果及分析；第 5 节是结语。

## 2. 基于机器学习的近义词辨析方法

### 2.1 训练集的准备

在近义词辨析任务中，机器学习模型的目标是：在给定的若干个近义词中，选出一个合适的词语填入句中特定位置，相比选择其他候选词填入该位置，能够使句子更为合理通顺。为此，机器需要比对一个候选词集中所有的候选词的上下文语境特征，找出候选词与其上下文语境之间的选择规律。由于候选的近义词相对于词汇全集来说是非常相近的（或者说是比较容易混淆的），仅仅将自然语料中正确句子的相关信息作为正例不足以让模型充分学习到近义词之间的语境区别。为了区分，需要人为构造可对比的正例和负例，成对地来学习句子表示（sentence representation）。例如：对候选词集“不管、尽管”和语料中的句子“哈雷 _____ 不富裕，还是买了车。”，分别将候选词填入句子，会形成一组训练样例：<正例：哈雷尽管不富裕，还是买了车。负例：哈雷不管不富裕，还是买了车。>。如果候选近义词集有超过两个候选项，同样可以此方式形成多组正例和负例。例如：对候选词集“二、两、俩”和语料中的句子“她升入了 _____ 班。”，会形成两组训练样例：<正例：她升入了二班。负例：她升入了两班。>；<正例：她升入了二班。负例：她升入了俩班。>。直观而言，在大多数情况下，将真实语料句子中的一个词随机替换成另一个词，得到的新的句子将是一个错误的句子，在统计意义上，以这样的方式自动构造的负例样本是可用的。

### 2.2 基于最大间距策略的训练过程

按照上述办法生成包含正例和负例的全部训练语料后，机器学习的过程就是训

练一个分类函数，让近义词辨析模型能最大程度区分填入正确候选词的句子与填入错误候选词的句子。具体方法是，采用 Bi-LSTM 对句子的信息进行捕捉，在观察大量的正例和负例句对过程中，学习一个分类函数，该函数为正例句子打出尽量接近 1 的分数，同时为负例句子打出尽量接近 0 的分数，使得正例句子的得分与负例句子的得分之差最大化[2]。在机器学习中，计算间距最大化（max-margin）的常用方法是采用折页损失函数（Hinge loss function）[3]。本文所用的函数形式为：**$L=max(0,1-(F_1-F_2))$**，其中 $F_1$ 代表正例句子的综合评分，$F_2$ 代表负例句子的综合评分。该损失函数的含义概括来说就是，如果填入候选词是正确的，则损失为 0，否则损失就是 1-（F1-F2）。为了使损失尽量小，就要求 F1 和 F2 的差距尽可能地大。

Bi-LSTM 的作用是训练一个词语用法模型（word usage model）。模型训练可以看做是记住了一个候选近义词的上下文特征，以及上下文中各个词语的权重，即上下文中各个词语对于选择某个特定候选词的影响大小。

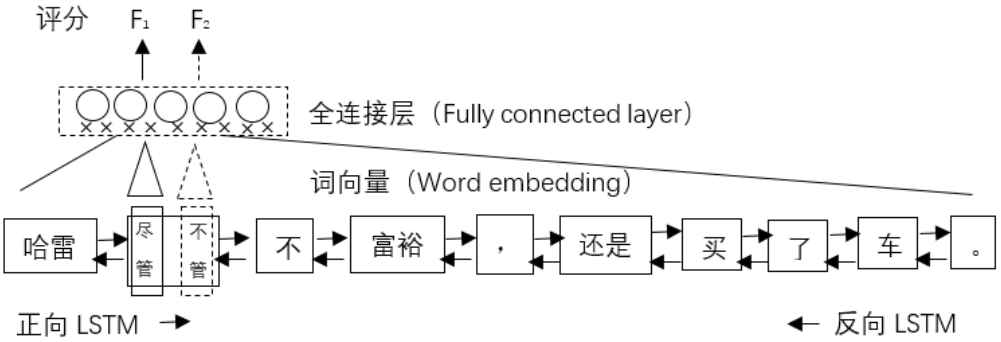



**图 1 “不管、尽管”辨析的 Bi-LSTM 机器学习模型示意[4]**

在图 1 所示的句子中填入“尽管”，为正例，计算综合评分 F1，填入“不管”，为负例，计算综合分值 F2，经过对包含“尽管”和“不管”的语料的充分训练后[5]，在图 1 所示的句例中，F1=0.9749742，F2=0.1040909。换言之，系统学习到一组参数，基于这些参数可计算得到“尽管”和“不管”在相同语境中的得分，且差距适当。

[2] 基于 Bi-LSTM 的机器学习模型在英语近义词辨析任务上取得了比较好的成绩，此外在构建基于文档的问答系统（Document-based Questions-Answer task，DBQA）等任务上也都有着不错的表现。

[3] 损失函数是机器学习训练参数的主要手段，即用损失函数来衡量模型学习到的参数的好坏。通常损失函数值为一个非负值。该值越小，表示模型的预测结果跟真实情况之间的差距越小，因而模型效果越好。

[4] Bi-LSTM 神经网络中，全连接层（Fully Connected Layer）也称为稠密层（Dense Layer）。该层的每一个节点都与前一层神经网络的每个节点连接，其作用简而言之就是把前一层学习到的分布式表示特征（比如一个句子的数百维向量表示）都综合起来，最终输出为一个值，这样便于完成分类任务。

[5] 在训练过程中，F1、F2 的值不断在变化。学习的过程就是通过 F1 和 F2 差值的变化，来调整句中词语的权重，从而找出最能影响正确选择候选词的上下文词语特征。比如机器通过观察大量语料，可能会发现规律：当上下文中出现了“还是”或者“不、没有”等否定词时，选择“尽管”的概率要高于选择“不管”。

### 2.3 对机器学习模型的测试

在测试阶段，将待测试的句子和候选近义词集作为机器学习系统的输入，机器学习模型就能够基于已有参数计算出候选近义词集中，哪一个更适合测试题的句子。仍以“尽管”和“不管”这组近义词为例，在本文实验中，测试题集第 70 题是“他没有来，_____ 我邀请了他。”，该题填入“尽管”后整句得分为 0.9999995，填入“不管”后整句得分为 0.0000134。机器学习系统将“尽管”作为答案输出。

## 3. 实验的准备

本文的主要目的是考察计算机辨析近义词的表现跟人类二语学习者的表现是否具有可比性，此外，通过比较机器学习系统在近义词辨析任务上的具体表现与人类二语学习者表现的相似或不同之处，探讨机器学习系统在二语教学中有可能发挥作用的途径。

因之前缺少在中文领域类似的实验研究可以借鉴，本文的工作属于初步尝试。在进行实验设计时，我们采取了在现有条件下尽可能平衡取样的做法，以便在控制实验成本的前提下尽量多地发现线索，为今后进一步深入研究打下基础。下面分别介绍近义词选词和测试题制作的情况，以及参与试验者样本情况和机器学习系统构建的情况。

### 3.1 近义词集的选样

前文提到的机器学习系统完成的近义词辨析实验往往只选取十组左右的近义词，而且通常仅限于实词性词语。这些研究聚焦在探讨机器学习系统的模型和算法，对近义词本身的语言学性质不太关注，此外，也较少从第二语言教学的角度去开展近义词辨析任务的相关研究。针对这些问题，本文在实验设计时主要考虑在以下三个方面有所加强：（1）要兼顾母语者视角和二语学习者视角的近义词，除一般的近义词（near-synonym words）外，还要选取部分易混淆词（confusable words）；（2）除一般实词外，还要选取一定比例的虚词；（3）尽量全面地考虑词语的频率、难度、义项数量等相关因素来平衡取样。

值得说明的是，关于易混淆词，对外汉语教学界有过不少探讨，特别是从二语学习者的视角来认识易混淆词，跟主要是从母语者视角认识的近义词（同义词）有很大的不同。张博（2007）指出，“‘易混淆词’与‘同义词’‘近义词’之间有交叉关系，而非包含关系或并列关系，因为它们是研究者站在不同的立场、以不同视角和不同标准归纳出来的词语类聚”。“易混淆词不仅体现为口头表达和写作中的词语混用，还体现为阅读和听辨中的词语误解。” 张博（2007）文中举出“往事-故事”“日前-目前”“有点儿-一点儿”“经验-经历”“乘（坐）-用”等多组易混淆词的例子，这些例子，都不是从母语者视角观察到的近义词，而是在对外汉语教学实践中发现的，由于各种原因（比如字形相近、母语影响等）造成的二语学习者容易混淆的词汇。

基于上述认识，本文选取了 67 个近义词（下文提及近义词时均含易混淆词，可认为是广义的近义词），共计 27 组（词集），作为辨析对象。下面表 1 和表 2 分别从词语类型分布和词集类型分布两个层面详细给出了本文所选词语和词集的统计信息。

**表 1 本文选取的 67 个近义词的类型分布**

<table>
<tr><td colspan="13">词语类型分布</td><td>合计</td></tr>
<tr><td rowspan="2">词类</td><td>动词</td><td>名词</td><td>形容词</td><td>副词</td><td>数词</td><td>数量词</td><td>量词</td><td>代词</td><td>介词</td><td>连词</td><td>助词</td><td>叹词</td><td rowspan="2">67</td></tr>
<tr><td>13</td><td>8</td><td>7</td><td>11</td><td>2</td><td>2</td><td>3</td><td>3</td><td>4</td><td>6</td><td>6</td><td>2</td></tr>
<tr><td rowspan="2">词汇等级1</td><td>HSK1</td><td colspan="2">HSK2</td><td colspan="2">HSK3</td><td>HSK4</td><td colspan="2">HSK5</td><td colspan="2">HSK6</td><td colspan="2">其他</td><td rowspan="2">67</td></tr>
<tr><td>13</td><td colspan="2">14</td><td colspan="2">7</td><td>17</td><td colspan="2">9</td><td colspan="2">1</td><td colspan="2">6</td></tr>
<tr><td rowspan="2">词汇等级2</td><td colspan="3">甲级</td><td colspan="3">乙级</td><td colspan="3">丙级</td><td colspan="4">丁级</td><td rowspan="2">67</td></tr>
<tr><td colspan="3">41</td><td colspan="3">14</td><td colspan="3">11</td><td colspan="4">1</td></tr>
<tr><td rowspan="2">义项数</td><td>单义</td><td>二义</td><td>三义</td><td>四义</td><td>五义</td><td>六义</td><td colspan="2">七义</td><td>八义</td><td rowspan="2">67</td></tr>
<tr><td>22</td><td>17</td><td>11</td><td>5</td><td>3</td><td>5</td><td colspan="2">3</td><td>1</td></tr>
</table>

**表 2 本文选取的 27 组近义词集的类型分布**

| 近义词辨析组的类型分布 | | | | 合计 |
|---|---|---|---|---|
| 义项关系 | 单义组 | 单义-多义组 | 多义组 | 27 |
| | 3 | 13 | 11 | |
| 词类关系 | 词类相同 | 词类有同有异 | 词类不同 | 27 |
| | 12 | 13 | 2 | |
| 词语数 | 二词辨析 | 三词辨析 | 四词辨析 | 27 |
| | 16 | 9 | 2 | |
| 语素 | 语素完全不同 | 有相同语素 | 语素逆序 | 27 |
| | 13 | 13 | 1 | |

附录 1 进一步给出了 27 组近义词集的描述信息，其中包括易混淆词集 11 组（40.74%）。另外值得说明的是，无论汉语本体研究还是对外汉语教学研究，因都是面向人的词汇教学，在考虑近义词辨析的因素时，通常都会兼顾多个层面的特征，包括句法特征、搭配、词语意义（含概念义、感情色彩义等）、语用特征（如语体风格、方言用法等）。在教学实践中，往往也强调从多个维度展开对具体近义词的辨析。不过，本文实验的测试对象包括计算机，在面向机器考虑近义词辨析因素时，只能分为两类，即有形式区别特征（包括广义句法形式特征和狭义的搭配词等）与无形

式区别的近义词[6]。附录 1 对此进行了大致的分类标记[7]，在 27 组近义词集中，无明显形式区别（即对人而言是意义差异）的近义词集有 13 组（48.15%）。

### 3.2 近义词辨析测试题集的制作

按照上述指导思想甄选出近义词集后，我们通过多种渠道来制作测试句集，包括在语料库、词典、相关研究文献中搜集合适的例句，以及自拟测试句，主要是设计辨析词项可以替换的语境和不可替换的语境。为尽可能真实地反映人的词汇知识和词汇运用能力，全部测试题均为不定项选择题，即除一般的单选题外，还有意设计了一定比例的多选题，并在所有备选答案之外增加了“我不知道”和“都不可以”选项（避免二语学习者猜答案）[8]。为了避免因词语偏难造成二语者理解上的障碍，测试句中的词语难度等级都控制在 HSK5 级以下，题干字数尽量控制在 20 字以内。然后再核查每个测试句的自然度、语境的适切度。在测试集第一版试题完成后，我们以在线问卷形式进行了调研，在收集了近 200 位母语者的语感调查数据后，我们增删了部分近义词集，并对试题进行了调整，最终制作了 100 道测试题的近义词辨析问卷。附录 2 给出了这 100 道题的题干和参考答案[9]，并对每道题根据是否有形式区别特征做了标注。表 3 是 100 道题的区别线索分类统计信息。其中“假搭配”特征是指题干中有意设计了某个候选近义词常见的共现形式特征，但整句语义排斥该候选词填入其中。

**表 3 近义词辨析测试集（100 题）根据是否具有形式区别线索分类的分布统计表**

| 特征类型 / 题型 | 无特征 | 搭配 | 句法形式 | 假搭配 | 合计 |
|---|---|---|---|---|---|
| 单选题 | 44 | 19 | 11 | 4 | **78** |
| 多选题 | 9 | 9 | 0 | 0 | **18** |
| 选“都不可以” | 1 | 2 | 0 | 1 | **4** |
| 合计 | **54** | **30** | **11** | **5** | **100** |

需要说明的是，机器学习系统要跟人类被试者一样完成这 100 道测试题。不过，

[6] 有意思的是，这里的所谓“无形式区别”又是从人的视角来说的，指没有“人通常理解的显式的语法形式区别”。按照人的语感，对这类近义词，一般是从意义的角度去把握其区别。但从计算机的视角来说，两个符号对象若存在区别，则一定是在形式上有区别。只不过，计算机察觉的“形式区别”，人不一定从“形式”的角度去看。对这类“形式区别”，通常笼统地概括为“意义的区别”。

[7] 这两类之间并无严格的界限。标注有一定的主观性。附录 1 中的分类标记是综合了多人标注的结果。

[8] 这些设置都是针对人类被试者的。目前的机器学习模型无法给出“我不知道”这个答案。而要针对多选和“都不可以”两种情况作答，需要对机器学习模型做重新设计，限于条件，本文实验所用的机器学习模型，仅实现了单选题的学习模型，无法像人那样回答答案是多选和“都不可以”这两种题型。

[9] 测试题的参考答案在进行评分时是作为标准答案看待的，即由程序自动比对被试答案与标准答案的异同来打分。不过，母语者语感调查显示，确实也有少数测试题的答案在母语者中也有语感差异，并不能做到完全一致。对于这些测试题，是按照少数服从多数原则，取多数母语者的答案作为参考答案。

机器学习系统是通过观察真实语料来学习近义词的语境分布差异，从而模拟人类的近义词辨析这一语言行为的。从统计学的角度讲，本文制作的 100 道题，并不是来自真实语料的样本，跟真实语料的分布有一定的差异，为了对比，同时也是为了更准确地反映机器学习系统在真实语料基础上所学习到的近义词辨析能力如何，我们在 100 题之外，还采用从真实语料中自动取样的大测试集对机器学习系统的近义词辨析能力进行了测试。参见下面 3.3 的具体说明。

### 3.3 被试者概况

实验的被试包括母语者（native speaker，记作 N）、二语者（second language speaker，记作 S），以及机器学习程序（Machine Learning System，记作 M）。其中二语者包括两组：一组是北京大学中文系的留学生（31 人），汉语水平达到 HSK6 级（记作 S_HSK6）；另一组是天津财经大学的留学生（28 人），汉语水平未达到 HSK6 级（记作 S_HSK6-）。他们来自韩国、日本、泰国、哈萨克斯坦、乌兹别克斯坦、赤道几内亚、加蓬、老挝、马来西亚、蒙古等 18 个国家。下面表 4（机器学习程序也当做一个被试看待）给出了本文被试的概况信息。

**表 4 母语者、二语者、机器学习程序的基本信息表**

| 组别 | 数量 | 年龄范围 | 平均年龄 | 学习时长 | 男 ： 女 |
|---|---|---|---|---|---|
| N | 20 | 21-56 | 29.85 | | 11 : 9 |
| S_HSK6 | 31 | 17-21 | 19.35 | 1.5 年-19 年 | 15 : 16 |
| S_HSK6- | 28 | 17-25 | 20 | 8 个月-2 年 | 12 : 16 |
| M | 3 | | | 126 小时 - 289 小时 | |

上表中 M 的数量标记为 3，指的是在程序开发和实验过程中，机器学习程序根据训练数据集大小不同有 3 个版本。机器学习程序的训练数据是从北京大学 CCL 语料库中抽取的包含 67 个近义词语的全部句子，经过自动分词和词性标注处理。训练集分为小、中、大三个版本，以观察数据规模变化对机器学习程序效果的影响。

**表 5 近义词辨析机器学习程序的训练数据集统计信息**

| 程序版本 | 训练数据集 | 字例数 | 字型数 | 词例数 | 词型数 |
|---|---|---|---|---|---|
| M30 | 30 万句训练集 | 28,237,314 | 6,475 | 15,239,701 | 184,985 |
| M100 | 100 万句训练集 | 89,674,369 | 7,286 | 48,760,493 | 293,659 |
| M300 | 300 万句训练集 | 228,081,378 | 8,199 | 122,791,248 | 466,176 |

我们也分别为这三个版本的机器学习程序构造了大测试集，分别为 6700 句，20000 句和 67000 句。三个版本在大测试集上的表现呈线性增长。以 67000 句测试集（即每个近义词对应的测试题为 1000 句）为例，M30、M100、M300 在该测试集上的正确率分别是 65.2%、65.9%、66.6%。在人工设计的 100 题测试集上，M30、M100、M300 的表现同样也呈线性增长关系，得分分别是 54、56、58。显然，表现

最好的版本是训练数据量最大的 M300。下文在介绍实验结果和进行分析讨论时，所用数据均为 M300 的数据。

## 4. 实验结果及分析

为实验操作方便，汉语母语者参与测试采用的是网上调查问卷的形式。二语学习者的测试则是课堂环境中进行的。在测试前一周发放近义词辨析词汇表（27 组 67 个词），学生可以查资料准备。在测试当天，学生拿到 100 题纸质测试卷后，基本在一节课时间（45 分钟）内完成答题（答题为闭卷形式，时间上无严格要求）。回收试卷后，在 Excel 表中对原始数据进行录入、核对，并跟计算机答题的结果一道，在程序辅助下，开展进一步的数据统计和分析。下面主要从三个方面，对实验结果进行概要介绍和分析。4.1 节说明测试成绩总体情况，在宏观层面分析各组被试者的基本特点；4.2 节考察各组数据间的差异显著性和相关系数，进一步分析机器学习系统的成绩跟二语学习者之间的可比性。4.3 节比较机器评测数据和二语学习者评测数据在各组近义词集上的成绩分布差异。

### 4.1 测试成绩总体情况

上文已经提到，实验被试包括 4 组。主要的实验是由 4 组被试完成 100 道测试题，此外，机器学习系统还在大测试集上进行了 27 组近义词的测试。因此，本文实验得到的数据可以分为 5 组，除下面表 6 中的 N、S_HSK6、S_HSK6-这三组外，机器学习程序的成绩有两组：机器学习程序大测试集成绩（下文以 M'标识）和机器学习程序 100 题成绩（以 M 标识）。另外，为更全面对照机器和二语者，有时也把 S_HSK6 和 S_HSK6-合并为 S。下面表 6 呈现了 N、S（S_HSK6、S_HSK6-）、M 在 100 题测试集上的总成绩对比情况。

**表 6 100 道测试题总成绩对比表（每题均计 1 分）**

| 题型 | 题量 | N | | S | | S_HSK6 | | S_HSK6- | | M | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 单选题 | 78 | 73.45 | 73.45 | 59.83 | 59.83 | 66.81 | 66.81 | 52.11 | 52.11 | 58 | 58 |
| 多选题 | 18 | 13.45 | 18 | 3.12 | 16.20 | 5.23 | 16.71 | 0.79 | 15.64 | 0 | 18 |
| 选“都不可以” | 4 | 3.45 | 3.45 | 1 | 1 | 1.71 | 1.71 | 0.21 | 0.21 | 0 | 0 |
| 总计 | 100 | 90.25 | 94.90 | **63.95** | **77.03** | 73.75 | 85.23 | 53.11 | 67.96 | **58** | **76** |

表 6 中每组被试的成绩都分为左右两列，左列是从严评分的分值，右列是从宽评分的分值。从严指的是被试选项跟标准答案完全一致才算对。从宽是针对多选题而言，只要被试选项是标准答案的子集，就算对。按照从严标准，M 的成绩介于 S_HSK6 和 S_HSK6-之间。如果仅从单选题成绩来看，M 的成绩跟 S 的单选题成绩非常接近（58:59.83）。按照从宽标准，M 的成绩也跟 S 的成绩非常接近（76:77.03）。

可以认为，本文实验的机器学习程序在近义词辨析任务上的表现，跟二语学习者应该是具有相当高的可比性的。

表 6 所展示的各组被试的总成绩分布也说明了本文对近义词的选取以及测试题的制作是比较合理的。在 100 题测试集上的表现，HSK6 级二语者全部超过 60 分，平均分比 HSK6 级以下二语者高 20 分。如果机器学习程序参与二语学习者成绩排名，在全部 59 名二语者中排第 43 名，在 28 名 HSK6 级以下二语者中排第 12 名（达到中等水平）。

此外值得一提的是，按从严评分标准，母语者单选题得分百分制为 94 分（73.45/78），多选题得分为 75；二语者这两项的得分分别是 77 和 17；机器这两项的得分[10]分别是 74 和 0。从多选题与单选题平均分的巨大差距可以看到，题型对成绩的影响非常显著。

将 100 题原始成绩按组汇总后可以得到 27 组近义词的成绩对照（附录 3）。其中机器在 100 题测试集上的成绩方差很大，不如在大测试集上的表现稳定（M 的方差是 M'的 3.5 倍），这也显示出 100 题的样本性质跟真实语料测试题之间的明显差异。

**表 7 六组测试成绩的均值、方差等基本统计量表**

| | M' | M | S_HSK6 | S_HSK6- | S | N |
|---|---|---|---|---|---|---|
| 平均值 | 0.662 | 0.608 | 0.751 | 0.558 | 0.659 | 0.905 |
| 方差 | 0.020 | 0.071 | 0.026 | 0.039 | 0.030 | 0.004 |
| 最高分 | 0.862 | 1.000 | 0.989 | 0.952 | 0.972 | 1.000 |
| 最低分 | 0.433 | 0.000 | 0.366 | 0.121 | 0.305 | 0.757 |
| 中值 | 0.665 | 0.667 | 0.763 | 0.518 | 0.648 | 0.920 |

母语者在 100 题上的成绩大致在 0.8 到 1 区间波动，方差仅为 0.004。显示母语者的近义词知识（语感）有很高共性，且表现稳定。各组数据按照降序排序后再加以对照，可以观察到不同被试组在近义词辨析表现上的稳定性有明显的差异（图 2）。从下降趋势线的斜率看，M'的数据最接近母语者，显示机器学习模型在大数据集上的训练和测试也具有较高的稳定性，但总体知识水平显著低于母语者，介于高级和中级二语学习者之间。二语者的下降斜率较大。而且高级水平二语者和中级水平二语者仅在中高分区域斜率相当，到低分区域，中级水平二语者的下降幅度显著高于高级水平二语者，显示在更为困难的近义词辨析任务上，两组二语者的语言能力差

[10] 本文的机器学习模型仅针对单选题进行训练，未考虑多选情况。因而对多选题也仅能当作单选题作答，包括零选题（答案为"都不可以"），也会选一个词作为答案输出。如按从宽评分标准，机器给出的多选题的答案，均在正确答案范围中，比二语者表现略好（机器与两组二语者的得分分别是 100,、90、87）。当然，18 道多选题中 2 选 2 的题有 8 道，3 选 3 的题 1 道（这意味着从宽评分标准下，可以白拿 9 分），3 选 2 的题 6 道，4 选 2 的题 3 道（后面这 9 道题使二语者和机器的得分稍微有了区别），从机器答题只选 1 个答案的角度看，机器按从宽标准计分也略占便宜（虽不容易选对，但也更不容易犯错）。

异较大。M 的下降趋势成阶梯状。这主要是因为机器在 100 题测试集上只是一个样本的成绩，跟其他组的数据为多个样本的平均分显著不同。从这点上说，可能 M’的成绩数据跟 S 更具可比性。

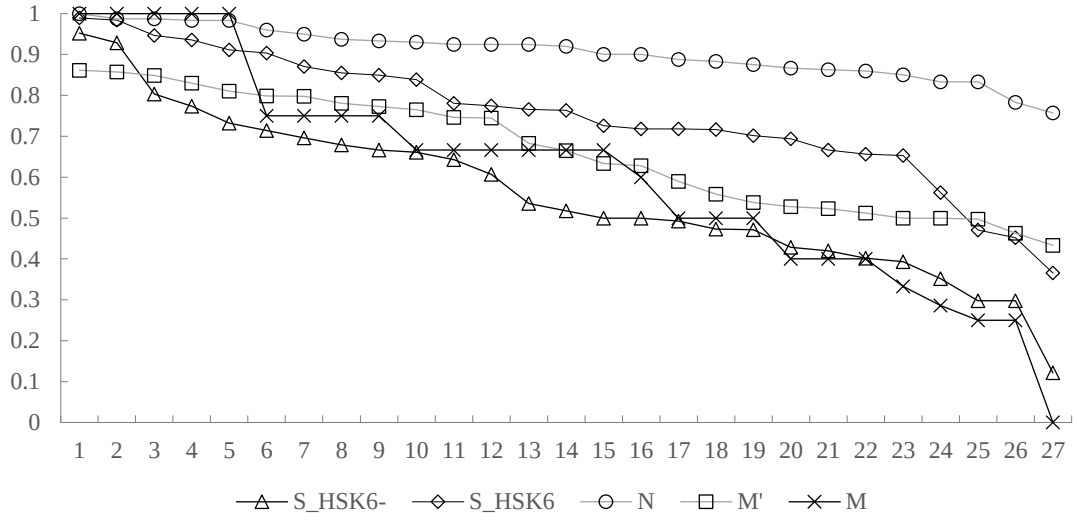



**图 2　在 27 组近义词上测试得到的五组成绩数据按降序排序**

## 4.2 各组测试成绩差异性与相关性分析

表 7 显示的机器测试成绩均值与二语者均值相对比较接近。但机器与二语者在本文测试集上的表现差异性到底如何，还需进一步进行统计检验。我们利用 Excel 内置的 T 检验功能（假设双样本异方差）[11]，对 6 组测试成绩数据进行了差异显著性分析，显著性水平 P 值设置为 0.05，结果如表 8 所示（表中数据为双尾 P 值，若 P>0.05，则表示差异不显著）：

**表 8 六组测试成绩数据间差异显著性分析结果**

| | M’ | M | N | S-HSK6 | S-HSK6- |
|---|---|---|---|---|---|
| M | **0.36** | | | | |
| N | 9.02E-10 | 4.18E-06 | | | |
| S-HSK6 | 0.035 | 0.022 | 0.00004 | | |
| S-HSK6- | 0.03 | **0.43** | 6.91E-10 | 0.00024 | |
| S | **0.95** | **0.41** | 5.55E-08 | - | - |

从表 8 检验结果可知，统计意义上，M’和 M 无显著差异，M’、M 分别和 S 无显著差异，M 和 S-HSK6-无显著差异（表中 P 值大于 0.05，以黑体标识）。从这个结果可以推测，机器在大规模真实语料上的近义词辨析测试结果可以在相当程度上

[11] 我们在 SPSS22.0 中使用 Kolmogorov-Smirnov 检验方法，发现各组样本数据均呈正态分布。在 Levene 方差齐性检验中，发现这六组数据的方差不相等。

预测二语者某个特定近义词辨析词集上的表现，特别是中级（HSK6 级以下）二语学习者的水平。

在考察了各组被试成绩的差异性之后，本文又进一步考察了各组被试在 27 组近义词集上的测试成绩的相关性。利用 Excel 内置的 CORREL 函数计算 6 组数据的线性相关系数结果如下面表 9 所示：

**表 9 六组测试成绩数据间相关系数**

| | M’ | M | N | S-HSK6 | S-HSK6- |
|---|---|---|---|---|---|
| M | 0.2402 | | | | |
| N | 0.2751 | 0.4085 | | | |
| S-HSK6 | 0.1520 | 0.6256 | 0.7449 | | |
| S-HSK6- | 0.1848 | **0.6810** | 0.5743 | 0.8886 | |
| S | 0.1742 | **0.6737** | 0.6741 | - | - |

结果显示，M 跟 S 两组数据的相关系数为 0.6737，经 T 检验公式计算（置信区间水平设为P=0.95）可知，M 跟 S 两组数据的 T 统计值为 5.8974，大于临界值 1.7081，因此，两组数据之间有较强的正相关。这在很大程度上说明，用机器学习系统进行近义词辨析测试所得的成绩数据，对于预估特定的近义词集和相应的一套测试题对于二语学习者的难度，有一定的参考价值。限于篇幅和数据处理成本方面的考虑，本文未进一步对近义词集分小类计算成绩相关系数，仅给出了总体相关性计算结果。如果分小类后再计算相关度，在某些具体类别的近义词集辨析任务上，机器学习系统的表现跟二语者的表现，可能会有更强的相关性。若据此为计算机辅助组织和实施近义词辨析的教学实践，应具有较高的参考价值。

### 4.3 机器与二语者答题分布差异考察与初步分析

近义词辨析，从名称来说，辨析的是词义的区别。就人的感觉而言，词义差别有的在用法上有形式区别，有的则仅有意义差别，而没有明显的形式区别。也正是基于这种语感，本文将 100 道测试题按照有无形式特征线索做了区分标记[12]（附录 2），以便考察被试在具体测试题上的表现跟近义词辨析形式线索之间的关系。我们统计了 100 题中得分大于 0.6 的题的形式特征标记分布情况，结果如下面表 10 所示。

**表 10 从形式特征角度看各组被试的答题正确率分布**

| 分值≥0.6 题数 | N | | S | | M | | 总题数 |
|---|---|---|---|---|---|---|---|
| 总题数 | 98 | 98.00% | 68 | 68.00% | 58 | 58.00% | 100 |
| 无形式特征题数 | 53 | 98.15% | 39 | 72.22% | 37 | 68.52% | 54 |

[12] 本文所说的形式特征包括句法形式特征、搭配（一般是特定词语）特征。无形式特征指从人的角度来看，只能根据对整句意思的理解来从候选词集中选择合适词语填入。对每道测试题做特征标注时，只能从“句法特征、搭配特征、无特征、假特征”四项中选择其一进行标注。

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 搭配特征题数 | 29 | 96.67% | 15 | 50.00% | 11 | 36.67% | 30 |
| 句法特征题数 | 11 | 100% | 11 | 100% | 8 | 72.73% | 11 |
| 含假搭配特征题数 | 5 | 100% | 3 | 60% | 2 | 40% | 5 |

需要说明的是，机器每道题的分值要么是 1（表示对），要么是 0（错），表 10 中 M 一列跟 N 和 S 两列的性质有所不同。N 和 S 两列的统计结果是基于所有参加测试的母语者和二语者在每道题上的平均分值。此外，上表是以全部 100 题作为统计对象的结果。其中二语者、机器的数据实际上仅限于单选题（二语者和机器多选题的分值均低于 0.6）。如前所述，二语者和机器在多选题上的正确率显著低于单选题，也远低于母语者正确率。对于母语者来说，如果以单选题作为统计范围，则各类题型的正确率均为 100%[13]。

从表 10 所示统计结果来看，人类二语者被试与机器学习系统对比有几个特点值得注意：（1）机器对“意义”的理解能力超出我们的预期。M 和 S 在“无形式特征”题中成绩超过 0.6 的题数数量接近，为 37:39。这意味着机器学习在没有明显形式特征提示线索的情况下，也能较好地模拟人的近义词辨析能力，做出比较恰当的选择。正如 Huang 等（2017）在英语近义词辨析任务实验时所发现的那样，Bi-LSTM 对句子之间的全局性差异有很好的把握能力，本文在汉语近义词辨析任务上的实验也展示了类似结果。基于神经网络的机器学习模型从句子中词语的分布表示（distributional representation）出发对句子建模，在一定程度上达到了从整体上把握句子“语义”的效果。（2）“搭配”特征在近义词辨析中通常作为主要手段进行分析和讲解，但在实际测试中，二语者和机器在有搭配线索题上的表现并不算太好。表 10 中的“搭配特征题”总数为 30，其中包含了多选题，如果去掉多选题，仅以 19 道单选题来说，在这些测试题上，得分超过 0.6 的比例，S 为 78.95%，M 为 57.89%。就机器来说，比例低于在无形式特征题上的表现。这提示我们，在统计意义上，搭配特征作为辨析线索，可能并不是以一种在统计上很凸显的模式出现，机器不见得能有效捕捉到[14]。（3）从“句法特征”与“搭配特征”的对比来看，机器在前一类测试题上的表现更好（72.73%>57.89%），这显示基于句法范畴的区别特征可能在分布上更具统计意义（或者说比“搭配特征”分布范围更广），机器更容易捕捉到这类信息。从二语者跟机器的成绩对比来看，二语学习者把握句法特征（或者说是“抽象形式”）的能力要强于机器。值得指出的是，含假搭配形式特征的题对于母语者没有起到“挖坑”效果，但对于二语者和机器，都造成了误判。不过由于样本量较小，不一定能说明问题，还有待今后设计更合理的实验来验证。

下面将附录 3 所示 27 组近义词测试成绩以折线图形式展示，以比较机器和二语者在具体的近义词集上的表现差异。

[13] 具体数据分别是：单选题总题数（78，100%），包括无形式特征题数（44，100%），搭配特征题数（19，100%），句法特征题数（11，100%），和假搭配特征题数（4，100%）

[14] 事实上，二语者的测试成绩也表明，学习者对搭配特征的把握效果并不理想。这说明从搭配的角度辨析近义词，对二语学习者也是比较大的挑战。

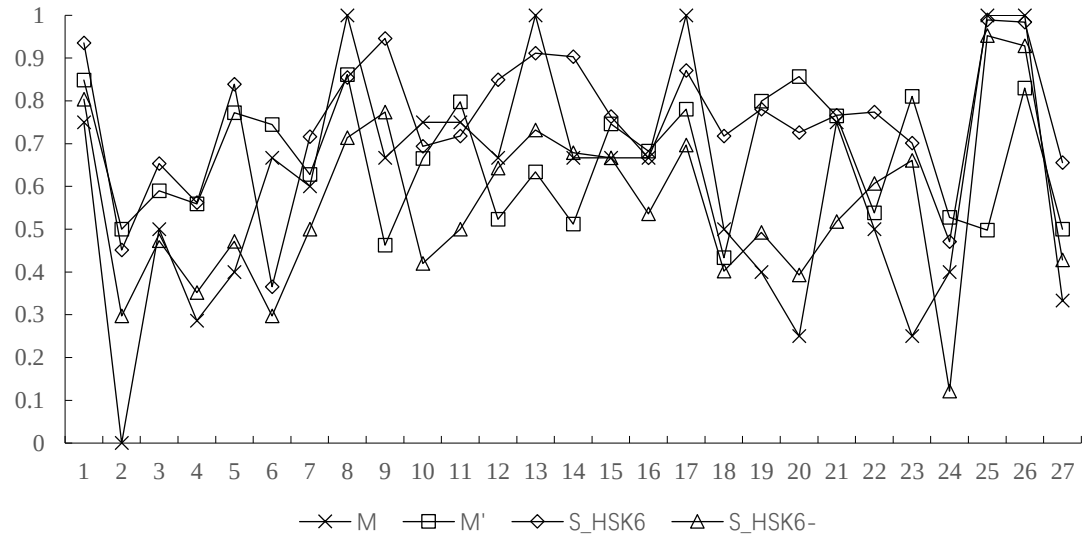



**图 3　在 27 组近义词上测试得到的四组成绩数据对照图**

整体而言，图 3 显示在大多数近义词集上，机器成绩与二语者成绩之间相差较大。像第 1 组这样成绩相对接近的近义词集占少数。就机器和二语者在每组具体的近义词集上的成绩表现而言，影响因素较多，较难概括出一般性的规律。比如像第 2、4、6、24 这四组近义词集，是二语者测试成绩排序最低的四组。机器成绩大多也较低，但跟二语者成绩并不相近。影响因素包括（1）近义词集中候选词较多（3 到 4 个）；（2）涉及题型包括多选题和答案为“都不可以”的题；（3）近义词本身辨析的形式线索较少或较难掌握，等等。机器成绩相对突出的有第 8、13、15、17、25、26 这几组，其中第 8 组“适合、合适”，第 15 组“有点儿、一点儿”，第 17 组“买、购买”，第 26 组“人、人们”的形式区别线索较为显著，机器和二语者得分均较高。这说明形式线索明确的近义词辨析任务相对容易完成。第 13（也、又、再、还）和 25（他、她、它）这两组，M 成绩很高，二语者成绩也相对较高，但 M’成绩却偏低，原因可能可以归结为人工测试题偏容易，但还有待进一步考察。类似的情形还出现在第 19（吗、吧、呢）、20（不管、尽管）、23（两、二、俩）这三组，M 跟 M’相差非常大，但关系倒过来，均为 M 远低于 M’。除 M 的题型难度更高（含多选、零选等）外，可能人工测试题跟真实训练语料的分布差异比较大也是影响因素，导致机器在真实语料测试集上表现正常，但在人工测试题上却容易出现误判。不过这也只是推测，有待进一步论证。

## 5. 结语

本文在前人有关近义词辨析的语言本体和教学研究基础上，将机器学习技术引入中文近义词辨析任务，对机器、母语者和二语学习者做了初步的实验对比研究。实验中所用近义词集的选择和试题的制作，均充分考虑平衡性和周全性原则。相比以往的类似实验，本文实验设计有两个特色：一是将机器测试与人类测试进行对照；二是将机器测试分为真实语料测试集和人工构造测试集进行对照。人工测试题的题

型设置比以往机器测试题的类型更丰富，更接近面向人类测试的题型设计。通过对测试成绩进行初步的统计和分析，我们得到了一些有参考意义的结论：（1）机器学习程序在中文近义词辨析任务上的总体表现，跟二语学习者具有明显的可比性。（2）机器学习程序跟二语学习者（特别是中级水平的汉语学习者）在近义词辨析任务上的总体表现有较强的相关性。（3）从二语者和机器学习系统在多选题与单选题上所得平均分的差距可以看到，题型对成绩的影响非常显著。（4）在影响近义词辨析的各项因素中，从“意义”（包括语体）角度辨析的重要性不亚于从形式区别特征的角度辨析，从抽象的句法形式特征角度进行辨析的重要性不亚于从搭配的角度进行辨析。

前人在英语近义词辨析任务上的实验表明，机器学习系统可以为二语者学习近义词提供辅助，为学习者自动推荐例句（Huang 等 2017）。本文在中文近义词辨析任务上的人机对照实验研究也有类似启示：机器学习技术有可能在汉语近义词辨析的教学方面起到辅助作用，可能的途径不仅包括推荐例句，还可以辅助教师构造近义词辨析测试题集，对测试集难度进行分级评估。此外，经过合理的系统设计，有可能形成人机互助的近义词辨析教学平台，由机器辅助人类教师构建近义词辨析知识库、题库，对近义词集的辨析难度，试题的难度进行分级，从而更为科学地组织和实施教学。从 NLP 的角度来说，目前近义词辨析任务无论是英文还是中文，都还缺乏广为接受的标准的可比测试集。本文在近义词集的选择和试题制作方面积累的经验，有可能为研制中文近义词辨析任务的标准测试集提供借鉴。

本文实验所用机器学习模型仅为基线系统（baseline system），即利用常用的深度学习技术构建的处理近义词辨析任务的程序。相比已有的在英文近义词辨析任务上的系统的表现来说，机器学习系统的设计还有优化空间，可以继续提升在中文近义词辨析任务上的性能表现。另外，在近义词集的选择、试题制作、试题类型标注等方面，此次试验均为全人工完成。未来可以更多地借助程序来进行辅助，包括内容设计和管理，针对某些特定类型的近义词集进行更为深入的专项测试等等，可以进一步提升实验的科学性和效率。

**致谢：**本文研究工作得到教育部人文社科重点研究基地重大项目（编号 13JJD740001、15JJD740002）和国家自然科学基金项目（编号 61876004）资助。北京大学中文系万艺玲老师和白一瑾老师为组织留学生参与此次实验提供了很多帮助。中文系现代汉语专业硕士研究生裴晓倩、施朝、柯丽珍等多人参与了母语者语感调查、近义词集筛选、测试题特征标注等工作，在此一并致谢！

## 参考文献

Chen, S., He, W., Prasetyo, P. K., & Yu, L. (2012, September). Skip N-gram Modeling for Near-Synonym Choice. In *Proceedings of the Twenty-Fourth Conference on Computational Linguistics and Speech Processing (ROCLING 2012)* (pp.163-175). [陈士婷,何维晟,关松坚,禹良治. (2012). 应用跳脱语言模型于同义词取代之研究.见第 24 届计算语言学和语音处理会议议程, 163-175.]

Edmonds, P. (1997, July). Choosing the word most typical in context using a lexical co-

occurrence network. In *Proceedings of the 35th annual meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics* (pp. 507–509). Madrid, Spain.

Gao, Z. (2016). A new probe into theory and method of comparing near-synonym: on teaching near-synonyms at these two levels. *Journal of Huaibei Normal University (Philosophy and Social Science)*, 2,141-145. [高再兰. (2016). 近义词比较理论与方法新探——兼谈两个层面的近义词教学.*淮北师范大学学报（哲学社会科学版）*, 2 , 141-145.]

Hong, J. (2014). Chinese Near-Synonym Study Based on the Chinese Gigaword Corpus and the Chinese Learner Corpus. In X. Su and T. He (Eds.). CLSW 2014, LNAI 8922 (pp. 329–340). Switzerland: Springer International Publishing.

Hong, W. (2012). A review of the synonyms studies in teaching Chinese as a second language. *TCSOL Studies*, 4, 44-51. [洪炜. (2012). 面向汉语二语教学的近义词研究综述.*华文教学与研究*, 4, 44-51.]

Hong, W. (2017). The effect of in-class differentiation on L2 acquisition of Chinese near-synonyms. *Chinese Language Learning*, 4, 84-91. [洪炜. (2017). 课堂显性辨析对汉语二语者习得近义词差异的影响.*汉语学习*, 4, 84-91.]

Huang, C., Chen, M., & Ku, L. (2017, April). Towards a Better Learning of Near-Synonyms: Automatically Suggesting Example Sentences via Filling in the Blank, *International World Wide Web Conference Committee* (IW3C2) (pp.293-302). Perth, Australia.

Lu, F. (2016). Corpus-based near-synonym teaching in TCSL. *Journal of Yunnan Normal University* (Teaching and Research on Chinese as a Foreign Language), 5, 49-56. [陆方喆. (2016). 基于语料库的对外汉语近义词教学.*云南师范大学学报（对外汉语教学与研究版）*, 5, 49-56. ]

McCarthy, D. & Navigli, R. (2007, June). The English lexical substitution task. In *Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007)* (pp. 48–53). Prague, Czech Republic.

McCarthy, D., & Navigli, R. (2009). The English lexical substitution task. *Language Resources and Evaluation*, 43(2), 139-159.

Nation, P. & Newton, J. (1997). Teaching vocabulary. In James Coady & Thomas Huckin. *Second Language Vocabulary Acquisition: A Rationale for Pedagogy* (pp.238-254). Cambridge: Cambridge University Press.

Wang, T. & Hirst, G. (2010, August). Near-synonym Lexical Choice in Latent Semantic Space. In *Proceedings of the 23rd International Conference on Computational Linguistics (Coling 2010)* (pp. 1182–1190). Beijing, China.

Yu, L., & Chien, W. (2013). Independent component analysis for near-synonym choice. *Decision Support Systems*, 55(1), 146-155.

Yu, L., Chien, W., & Chen, S. (2011, November). A Baseline System for Chinese Near-Synonym Choice. In *Proceedings of the 5th International Joint Conference on Natural Language Processing* (pp. 1366–1370). Chiang Mai, Thailand.

Zhang, B. (2007). Synonymy, near-synonymy and confusable word: A perspective transformation from Chinese to interlanguage. Chinese Teaching in the World,3,98-107. [张博. (2007). 同义词、近义词、易混淆词：从汉语到中介语的视角转移.*世界汉语教学*, 3, 98-107.]

Zhao, S., Zhao, L., Zhang, Y., Liu, T., &Li, S. (2007, June). HIT: Web based Scoring Method for English Lexical Substitution. In *Proceedings of the 4th International Workshop on Semantic Evaluations (SemEval-2007)* (pp.173–176). Prague, Czech Republic.

Zhao, X., & Liu, R. (2005). Some basic problems in compiling *Practical dictionary of near-synonyms for foreigners. Lexicographical Studies*,4,57-67. [赵新,刘若云. (2005). 编写《外国人实用近义词词典》的几个基本问题.*辞书研究*, 4, 57-67.]

**附录 1 本文实验所用的近义词（含易混淆词）组（27 组词）**

| 序号 | 词语 1 | 词语 2 | 词语 3 | 词语 4 | 近义/易混淆 | 形式区别特征 | | |
|---|---|---|---|---|---|---|---|---|
| | | | | | | 词类 | 句法形式 | 搭配 |
| 1 | 为了 | 因为 | | | 易混淆 | | | |
| 2 | 难过 | 难受 | | | 近义 | | | |
| 3 | 一再 | 再三 | | | 近义 | | ┬ | ┬ |
| 4 | 认识 | 了解 | 知道 | 理解 | 近义 | | | |
| 5 | 不 | 没 | | | 近义 | | | |
| 6 | 经验 | 经历 | | | 易混淆 | | | |
| 7 | 能 | 可以 | 会 | | 近义 | | | |
| 8 | 合适 | 适合 | | | 近义 | ┼ | ┼ | |
| 9 | 结果 | 后果 | 成果 | | 近义 | | | |
| 10 | 骄傲 | 自豪 | | | 近义 | | ┬ | |
| 11 | 可惜 | 遗憾 | | | 近义 | | ┬ | |
| 12 | 还是 | 要么 | 或者 | | 近义 | | ┬ | ┬ |
| 13 | 再 | 也 | 又 | 还 | 近义 | | | |
| 14 | 希望 | 愿望 | | | 近义 | ┬ | | |
| 15 | 有点儿 | 一点儿 | | | 易混淆 | ┼ | ┼ | |
| 16 | 相信 | 信任 | | | 近义 | | | |
| 17 | 买 | 购买 | | | 近义 | | ┬ | ┬ |
| 18 | 次 | 趟 | 遍 | | 易混淆 | | | |
| 19 | 不管 | 尽管 | | | 易混淆 | | ┬ | ┬ |
| 20 | 吗 | 吧 | 呢 | | 易混淆 | | ┬ | ┬ |
| 21 | 着 | 了 | 过 | | 易混淆 | | | ┬ |
| 22 | 哎 | 唉 | | | 易混淆 | | | |
| 23 | 二 | 两 | 俩 | | 近义 | ┬ | | |
| 24 | 向 | 往 | 朝 | | 易混淆 | | | |
| 25 | 他 | 她 | 它 | | 易混淆 | | | |
| 26 | 人 | 人们 | | | 易混淆 | | ┼ | |
| 27 | 千万 | 万万 | | | 近义 | | ┬ | |

说明：表中“形式区别特征”栏的单元格中填“┼”，表示有较明显的差异（比如“适合”跟“合适”分别是动词和形容词，句法差异较大）；填“┬”表示有一定的差异（比如“一再”跟“再三”在状语位置区别不明显，但“再三”可以出现在动词后，如“考虑再三”。“一再”没有这种用法）。不填表示无明显的形式区别线索。主要是意义上的区别。

**附录 2 近义词辨析测试题（27 组词语，共 100 题）**

| 题号 | 组号 | 题干 | 特征 | 参考答案 |
|---|---|---|---|---|
| 1 | 1 | 我们（　　）解决这个问题采取了不少的方法。 | 无特征 | 为了 |
| 2 | 1 | 她来中国学习汉语，是（　　）她觉得汉语很有用。 | 无特征 | 因为 |
| 3 | 1 | （　　）他一个人在国外学习，所以他的妈妈非常担心他。 | 搭配 | 因为 |
| 4 | 1 | 我们这次活动是（　　）保护环境，所以希望大家都能参与进来。 | 假搭配 | 为了 |
| 5 | 2 | 他们听说这件事以后，心里非常（　　）。 | 搭配 | 难过\|难受 |
| 6 | 2 | 他生病了，总咳嗽，非常（　　）。 | 搭配 | 难受 |
| 7 | 2 | 她的脸色非常（　　），像是在生气。 | 搭配 | 都不可以 |
| 8 | 3 | 他（　　）表示今后一定努力学习汉语。 | 无特征 | 一再\|再三 |
| 9 | 3 | 你的成绩为什么（　　）下降？ | 无特征 | 一再 |
| 10 | 3 | 他考虑（　　），最后决定开一家书店。 | 句法 | 再三 |
| 11 | 3 | 他（　　）再四地向老师表示感谢。 | 搭配 | 再三 |
| 12 | 4 | 要想真正地（　　）一个国家，最好去那个国家生活一段时间。 | 无特征 | 认识\|了解 |
| 13 | 4 | 我完全不能（　　）你现在的心情。 | 无特征 | 理解 |
| 14 | 4 | 他的家人都不（　　）他的病情。 | 无特征 | 了解\|知道 |
| 15 | 4 | 我（　　）小明什么时候回国。 | 无特征 | 知道 |
| 16 | 4 | 我们已经（　　）很多年了。 | 无特征 | 认识 |
| 17 | 4 | 我的家人都十分（　　）妈妈的脾气。 | 搭配 | 了解 |
| 18 | 4 | 员工们都十分（　　）公司目前的情况。 | 搭配 | 了解\|理解 |
| 19 | 5 | 小明为什么（　　）去爬长城呢？ | 无特征 | 不\|没 |
| 20 | 5 | 昨天下大雨了，所以我（　　）去爬长城。 | 无特征 | 没 |
| 21 | 5 | 在公共场所，请（　　）吸烟。 | 搭配 | 都不可以 |
| 22 | 5 | 明天我要去机场接朋友，所以（　　）去爬长城了。 | 句法 | 不 |
| 23 | 5 | 对不起，你说得太快了，我（　　）听懂你说的话。 | 句法 | 没 |
| 24 | 6 | 他的生活（　　）非常丰富。 | 搭配 | 经验\|经历 |

| | | | | |
|---|---|---|---|---|
| 25 | 6 | 她当了三十年的汉语老师，有着丰富的教学（　　）。 | 无特征 | 经验 |
| 26 | 6 | 他希望以后再也不要重复这段痛苦的（　　）了。 | 无特征 | 经历 |
| 27 | 7 | 现在我的牙不疼了，(　　)吃饭了。 | 无特征 | 能\|可以 |
| 28 | 7 | 他特别（　　）睡，一睡就是一整天。 | 无特征 | 能 |
| 29 | 7 | 你有空儿的时候，（　　）看看电影放松放松。 | 无特征 | 可以 |
| 30 | 7 | 你们不要再吵了，这样下去我（　　）失眠的。 | 无特征 | 会 |
| 31 | 7 | 让他一个人出差（　　）行吗？ | 搭配 | 能 |
| 32 | 8 | 这件西装不大不小，正（　　）。 | 句法 | 合适 |
| 33 | 8 | 她非常喜欢吃辣的，四川菜正（　　）她的口味。 | 句法 | 适合 |
| 34 | 9 | 比赛刚结束，运动员都在等待比赛（　　）。 | 搭配 | 结果 |
| 35 | 9 | 根据最新研究，地下水一旦被污染，将会产生极其严重的（　　）。 | 搭配 | 后果 |
| 36 | 9 | 这项科技（　　）将运用到人们的生活中。 | 搭配 | 成果 |
| 37 | 10 | 我们为自己的祖国感到（　　）！ | 搭配 | 骄傲\|自豪 |
| 38 | 10 | 虽然他得了好几次第一名，但他一点儿也不（　　）。 | 无特征 | 骄傲 |
| 39 | 10 | 这个机会增强了她的自信心和（　　）感。 | 无特征 | 自豪 |
| 40 | 10 | 蜿蜒万里的长城是中华民族的（　　）。 | 无特征 | 骄傲 |
| 41 | 11 | 这么好的机会，他都错过了，挺（　　）的！ | 无特征 | 可惜\|遗憾 |
| 42 | 11 | 对于这一事件的发生，我们深表（　　）。 | 搭配 | 遗憾 |
| 43 | 11 | 这件衣服还能穿，扔了有点儿（　　）。 | 无特征 | 可惜 |
| 44 | 11 | 这是一次非常难得的机会，（　　）他没有抓住。 | 无特征 | 可惜 |
| 45 | 12 | 你想喝茶（　　）喝咖啡？ | 无特征 | 还是 |
| 46 | 12 | 你要么战胜困难，（　　）被困难战胜。 | 搭配 | 要么 |
| 47 | 12 | 我们坐地铁去（　　）打车去都可以。 | 无特征 | 或者 |
| 48 | 13 | 你刚才唱得真好听，（　　）唱一个吧。 | 无特征 | 再 |
| 49 | 13 | 这是老王的主意，我（　　）同意。 | 无特征 | 也 |

| | | | | |
|---|---|---|---|---|
| 50 | 13 | 你每年生日的时候，我都会送你一件生日礼物。明天（　　）是你的生日了，你想要什么生日礼物呢？ | 无特征 | 又 |
| 51 | 13 | 我记得你十几年前在中国学过汉语，现在（　　）会说吗？ | 无特征 | 还 |
| 52 | 14 | 我对未来充满了（　　）和信心。 | 无特征 | 希望 |
| 53 | 14 | （　　）我们这次合作愉快。 | 无特征 | 希望 |
| 54 | 14 | 他通过努力终于实现了自己心中的（　　）。 | 无特征 | 愿望 |
| 55 | 15 | 他现在（　　） 累，想休息一会儿。 | 句法 | 有点儿 |
| 56 | 15 | 我觉得今年夏天比去年热（　　）。 | 句法 | 一点儿 |
| 57 | 15 | 这件衣服的质量（　　）好。 | 无特征 | 都不可以 |
| 58 | 16 | 我（　　）你一定能学好汉语。 | 无特征 | 相信 |
| 59 | 16 | 夫妻之间应该互相关心，互相（　　）。 | 搭配 | 信任 |
| 60 | 16 | 公司的老板非常（　　）他。 | 无特征 | 相信\|信任 |
| 61 | 17 | 太饿了，我去超市（　　）个面包吃。 | 句法 | 买 |
| 62 | 17 | 本店新到一批服装，欢迎新老顾客前来（　　）。 | 无特征 | 购买 |
| 63 | 17 | 如果你同意（　　）房的话，我们可以去售楼处看看。 | 无特征 | 买 |
| 64 | 17 | 双方当场就签订了（　　）合同。 | 无特征 | 购买 |
| 65 | 18 | 这个月他去了三（　　）广州了。 | 无特征 | 次\|趟 |
| 66 | 18 | 在这半年里，他先后动了三（　）手术了。 | 无特征 | 次 |
| 67 | 18 | 他花了十年的时间把图书馆里的书都读了一（　）。 | 无特征 | 遍 |
| 68 | 18 | 这（　　）开往北京的列车就要出发了。 | 搭配 | 趟 |
| 69 | 19 | （　　） 遇到的困难大不大，我们都要完成这项任务。 | 句法 | 不管 |
| 70 | 19 | 他没有来，（　　）我邀请了他。 | 无特征 | 尽管 |
| 71 | 19 | （　　）天冷天热，他一运动就会出很多汗。 | 句法 | 不管 |
| 72 | 19 | （　　）怎么解决好大家的困难，是他最关心的。 | 假搭配 | 都不可以 |
| 73 | 19 | （　　）遇到的困难很大，但我们也要坚持下去。 | 句法 | 尽管 |

| | | | | |
|---|---|---|---|---|
| 74 | 20 | 你是新来的同学（　　）？ | 搭配 | 吗\|吧 |
| 75 | 20 | 要是我不说，难道你就不认识我了（　　）？ | 假搭配 | 吗 |
| 76 | 20 | 我们一直在这儿等他，要是他不来了（　　）？ | 搭配 | 呢 |
| 77 | 20 | 我大概拿错书了（　　）？这好像不是我的书。 | 假搭配 | 吧 |
| 78 | 21 | 他看（　）半天菜单，一个菜都没有点。 | 无特征 | 了 |
| 79 | 21 | 爷爷喜欢在沙发上坐（　　）看报纸。 | 无特征 | 着 |
| 80 | 21 | 我从来没有在大草原上骑（　　）马。 | 假搭配 | 过 |
| 81 | 21 | 他关注和思考的是随着历史发展而不断变化（　　）的人和人的关系。 | 无特征 | 着 |
| 82 | 22 | （　　），我倒是有一个好主意。 | 无特征 | 哎 |
| 83 | 22 | （　　），他长长地叹了一口气。 | 无特征 | 唉 |
| 84 | 23 | 开学了，玛丽刚升到（　　）班学习汉语。 | 搭配 | 二 |
| 85 | 23 | 医生让他在家里好好儿休息（　　）天。 | 搭配 | 两 |
| 86 | 23 | 他有（　　）朋友，他们的关系非常好。 | 无特征 | 俩 |
| 87 | 23 | 老板，给我来（　　）斤苹果。 | 搭配 | 二\|两 |
| 88 | 24 | 年轻人，你们正在走（　　）美好的未来，祝福你们！ | 搭配 | 向 |
| 89 | 24 | 他说的这些话都是气话，你可别（　　）心里去。 | 搭配 | 往 |
| 90 | 24 | 这位明星（　　）我们挥了挥手。 | 搭配 | 朝\|向 |
| 91 | 24 | 这条小路通（　　）山顶。 | 搭配 | 向\|往 |
| 92 | 24 | 小孩子们（　　）东边跑去了。 | 无特征 | 朝\|向\|往 |
| 93 | 25 | 这位姑娘的爸爸是一名医生，（　　）的工作很辛苦。 | 无特征 | 他 |
| 94 | 25 | 这位男护士的女朋友长得非常漂亮，可最近听说这位男护士和（　　）分手了。 | 无特征 | 她 |
| 95 | 25 | 他奶奶家里养着一条狗，（　）的毛是棕色的。 | 无特征 | 它 |
| 96 | 26 | 傍晚，很多（　　）在公园里散步。 | 无特征 | 人 |
| 97 | 26 | （　　）都说他是一位好医生。 | 无特征 | 人们 |
| 98 | 27 | 他（　　）没有想到，苦心经营了五年的公司就这样倒闭了。 | 搭配 | 万万 |

| | | | | |
|---|---|---|---|---|
| 99 | 27 | 你们在外面（　　）要小心，不要被骗了。 | 搭配 | 千万 |
| 100 | 27 | 这件事情特别重要，（　　）不可马虎。 | 搭配 | 万万\|千万 |

问卷样题：
请在划线处填入合适的词语，可以填一个或多个。如果没有合适的词语可填，就填“×”（表示“都不可以”），如果不知道该选哪个词，就填“我不知道”。例如：

A 时候　　B 时间

（1）请不要忘记我们集合的__B__和地点。
（2）__A 、B__不早了，我该走了。
（3）刚到中国__×__，我一句汉语都不会说。
（4）初学汉语__我不知道__，我感觉很迷茫。

**附录 3 近义词辨析实验成绩列表（成绩取值在 0-1 之间）**

| 组号 | 近义词项 | M’ | M | S_HSK6 | S_HSK6- | S | N |
|---|---|---|---|---|---|---|---|
| 1 | 为了 因为 | 0.849 | 0.750 | 0.935 | 0.804 | 0.873 | 0.988 |
| 2 | 难受 难过 | 0.500 | 0.000 | 0.452 | 0.298 | 0.379 | 0.883 |
| 3 | 一再 再三 | 0.590 | 0.500 | 0.653 | 0.473 | 0.568 | 0.863 |
| 4 | 了解 理解 知道 认识 | 0.559 | 0.286 | 0.562 | 0.352 | 0.462 | 0.757 |
| 5 | 不 没 | 0.773 | 0.400 | 0.839 | 0.471 | 0.664 | 0.930 |
| 6 | 经历 经验 | 0.745 | 0.667 | 0.366 | 0.298 | 0.333 | 0.783 |
| 7 | 会 可以 能 | 0.628 | 0.600 | 0.716 | 0.500 | 0.614 | 0.920 |
| 8 | 合适 适合 | 0.862 | 1.000 | 0.855 | 0.714 | 0.788 | 0.925 |
| 9 | 后果 成果 结果 | 0.463 | 0.667 | 0.946 | 0.774 | 0.864 | 0.933 |
| 10 | 自豪 骄傲 | 0.665 | 0.750 | 0.694 | 0.420 | 0.564 | 0.888 |
| 11 | 可惜 遗憾 | 0.798 | 0.750 | 0.718 | 0.500 | 0.614 | 0.938 |
| 12 | 或者 要么 还是 | 0.523 | 0.667 | 0.849 | 0.643 | 0.751 | 0.900 |
| 13 | 也 再 又 还 | 0.634 | 1.000 | 0.911 | 0.732 | 0.826 | 0.925 |
| 14 | 希望 愿望 | 0.512 | 0.667 | 0.903 | 0.679 | 0.797 | 0.983 |
| 15 | 一点儿 有点儿 | 0.747 | 0.667 | 0.763 | 0.667 | 0.718 | 0.867 |
| 16 | 信任 相信 | 0.683 | 0.667 | 0.667 | 0.536 | 0.605 | 0.833 |
| 17 | 买 购买 | 0.781 | 1.000 | 0.871 | 0.696 | 0.788 | 1.000 |
| 18 | 次 趟 遍 | 0.433 | 0.500 | 0.718 | 0.402 | 0.568 | 0.900 |
| 19 | 不管 尽管 | 0.799 | 0.400 | 0.781 | 0.493 | 0.644 | 0.960 |
| 20 | 吗 吧 呢 | 0.857 | 0.250 | 0.726 | 0.393 | 0.568 | 0.925 |
| 21 | 了 着 过 | 0.765 | 0.750 | 0.766 | 0.518 | 0.648 | 0.988 |
| 22 | 哎 唉 | 0.538 | 0.500 | 0.774 | 0.607 | 0.695 | 0.850 |
| 23 | 两 二 俩 | 0.811 | 0.250 | 0.702 | 0.661 | 0.682 | 0.875 |
| 24 | 向 往 朝 | 0.528 | 0.400 | 0.471 | 0.121 | 0.305 | 0.860 |
| 25 | 他 她 它 | 0.498 | 1.000 | 0.989 | 0.952 | 0.972 | 0.983 |
| 26 | 人 人们 | 0.830 | 1.000 | 0.984 | 0.929 | 0.958 | 0.950 |
| 27 | 万万 千万 | 0.500 | 0.333 | 0.656 | 0.429 | 0.548 | 0.833 |

http://www.tclt.us/journal/2019v10n1/xulili.pdf pp. 26-56

# Using Online Applications to Improve Tone Perception among L2 Learners of Chinese
# (网络应用对中文二语学习者声调辨识的有效性研究)

Xu, Hongying
(徐红英)
University of Wisconsin-La Crosse
(威斯康星大学拉克罗校区)
hxu@uwlax.edu

Li, Yan
(李艳)
University of Kansas
(堪萨斯大学)
liyan@ku.edu

Li, Yingjie
(李颖颉)
University of Colorado-Boulder
(科罗拉多大学博尔得分校)
Yingjie.Li@colorado.edu

**Abstract**: This study investigated the effectiveness of an online application in helping beginning-level Chinese learners improve their perception of the tones in Mandarin Chinese. Two groups—one experimental and one traditional—of beginning Chinese learners from two universities in the Midwest participated in this study. The experimental group (n=20) used the online application to practice tones for four 15-minute sessions in class. The traditional group (n=11) participated in traditional instructor-led practice in class in lieu of the online practice. Both groups completed a pre-test, an immediately administered post-test, and a delayed post-test designed to assess their perception of the tones of monosyllabic and disyllabic words. No statistically significant difference has been found between the two groups in their tone perception performance in the post-test and in the delayed post-test. However, the experimental group showed a positive trend in improving their perception on those tones which posed more difficulty than others. Their experience with this online application and the pronunciation learning strategies of participants in the experimental group were also examined through a survey. Based on the findings, it is proposed that the use of online tone practice is worthwhile in a Chinese language class, but might fit better into the curriculum as external assignments.

**摘要：**本研究调查了网络应用在帮助初级中文二语学习者提高声调辨识的准确度上的有效性。实验组和对照组的参与者均来自于美国中西部两所四年制大学的初级中文班。实验组的学生（n=20）参与了四次每次十五分钟的基于网络应用的课堂练习，对照组的学生（n=11) 则在实验组使用网络应用进行声调练习的时间内进行传统的教师带领的声调练习。两组学生都参加了测试单音节词、双音节词中的声调辨识准确度的前测、后测和延时后测。实验组与对照组在前测、后测及延后测试中声调辨识的准确度上没有显著差异。然而，实验组在某些比较难辨识的声调上显示出较好的改进倾向。实验组的学生还完成了一份检测学习者使用网络应用进行声调练习的经验和学习发音

的策略的问卷。根据调查结果，本文建议在正式的教学中使用网络应用进行声调练习，但这种练习可能更适合作为作业让学生在课外完成。

**Keywords**: Online application, Chinese tone perception, learning experience

**关键词**：网络应用、中文声调辨识、学习体验

## 1. Introduction

As a tonal language, Mandarin Chinese employs different pitch heights and pitch contours to distinguish meaning. Research has shown that, regardless of their first language (L1), second language (L2) learners of Chinese have difficulty mastering Mandarin tones (Miracle, 1989; Shen, 1989; Shen & Lin, 1991; Sun, 1998; Jongman, Wang, Moore, & Sereno, 2006; Lee, Tao, & Bond, 2010; Chun, Jiang, Meyr, & Yang, 2015; Hao, 2012).

Despite the fact that Chinese language instructors and Chinese linguists recognize the critical role that the mastery of tones plays in developing communicative competence, the teaching of tones in the classroom lags far behind the teaching of other linguistic aspects and skills (Xing, 2006; Orton, 2013).

A key to channeling the genuine need and practice is finding an efficient way to improve learners' perception of tones in a short period of time. Incorporating online tone training practice into Chinese language class might solve the dilemma. Some studies carried out in lab settings show that intensive tone training can improve Chinese learners' perception accuracy on tones (Wang, Spence, Jongman, & Sereno, 1999) which can, in turn, be transferred to tone production (Rochet 1995; Wang, Jongman, & Sereno, 2003). Would similar effects be observed among Chinese learners who receive online tone training in a classroom setting? How would students of Chinese feel about online tone training? These questions remain unanswered because hardly any empirical studies that test the effectiveness of online tone training sessions on learners' perception of tones have been carried out in authentic classroom situations. In addition, very few studies have investigated the training effects of tone perception using true beginners during the time of their initial exposure to tones. Studies of the training effect in regular classrooms with true L2 Chinese beginners are thus needed.

The present study addresses this gap by exploring the effectiveness of computer-assisted online tone practice incorporated into classroom instruction in improving true beginning L2 Chinese learners' perception of tones. In addition, this study also investigates L2 Chinese learners' experience in their use and perception of online tone practice.

This article is organized as follows: Section 2 discusses research on L2 Chinese learners' tonal perception and the literature on computer-assisted tone training; Section 3 reports on the current empirical study, including research questions, experimental design, results and discussion; Section 4 comprises the conclusion drawn from the experiment and

the pedagogical implications, and; Section 5 discusses the limitations of the current study and suggests directions for future research.

## 2. Overview of L2 Chinese Learners' Tonal Perception and Current Literature on Computer-Assisted Tone Training

### 2.1 L2 Chinese Learners' Tonal Perception

Tones in Chinese are lexical in that the application of different tones to the same syllable results in different meanings. Linguists generally agree that there are four tones plus a neutral tone in Mandarin Chinese. These tones are different in pitch height and contour as can be seen in Fig. 1 below, in which a scale of 1 to 5 is used to describe pitch height and the critical points in the change of pitch height, i.e. the beginning point and endpoint, are depicted to represent different tones.

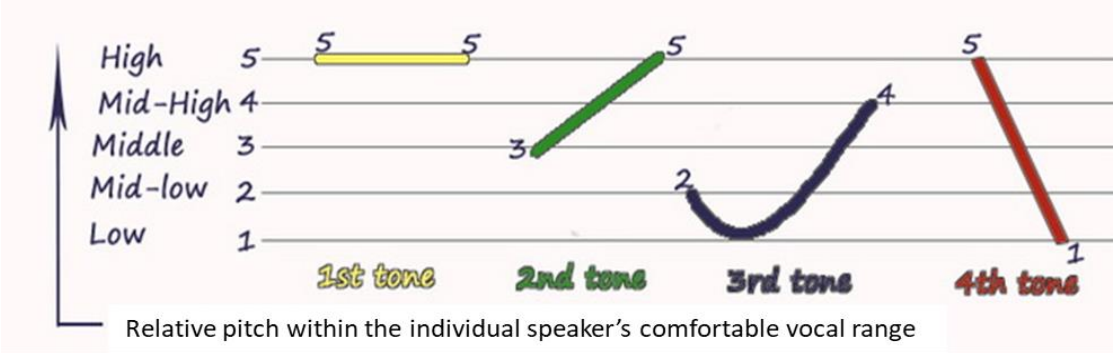



**Figure 1 Tone Pitch and Contour Chart**

The same Chinese syllable can express several unrelated meanings when pronounced using different tones. For example, if the syllable *ma* is pronounced using the first tone, a high level tone represented by 55, it means "mom"; when it is pronounced using the second tone, 35, it means "hemp"; when it is pronounced using the third tone, 214, it means "horse"; and it means "to scold" when pronounced using the fourth tone, 51. The neutral tone differs from the four tones in that its pitch height is decided by the tone of the preceding syllable and its pronunciation is light and short. The syllable *ma*, when pronounced using the neutral tone, functions as a question particle.

A failure to perceive Mandarin tones and establish the tonal categories can "have direct and drastic consequences" (Rohr, 2014) on both listening comprehension and communicative competence. It has been widely reported that L2 Chinese learners whose L1 has no tonal element have difficulty perceiving tones in Mandarin (Miracle, 1989; Shen, 1989; Shen & Lin, 1991; Sun, 1998; Jongman et al., 2006; Lee, et al., 2010). On monosyllabic words, L2 Chinese learners seem to acquire the rising tone (Tone 2) or dipping tone (Tone 3) later than the level tone (Tone 1) or the falling tone (Tone 4). A few studies that examined learners' tonal perception on disyllabic words found that learners' tonal performance on disyllabic words is generally worse than on monosyllabic words due to the complex relationship between adjacent tones and the great tonal variability in disyllabic words.

## 2.2 Research on Computer-assisted Tone Training

In today's Chinese language classroom in the United States, the teaching of pronunciation, including tones, lags far behind the teaching of other linguistic aspects and skills (Levis, 2007). Tones are introduced to learners mainly in isolation within a short period at the very beginning of the target language instruction: no priority is assigned to instruction of tones and less time is allotted to practice (Xing, 2006; Orton, 2013).

Various factors contribute to this issue, some of which are quite justifiable. Students exhibit wide-ranging differences in pronunciation needs and challenges: some need help with retroflex fricatives; some need to work on the palatals; some want more practice with third tones, etc. It is almost impossible to cater to all individual needs in class. Some language instructors, without foundational training in phonetics and phonology, have a hard time effectively explaining the articulation position and manner of each sound or each tone. Thus, when their students make mistakes, it is hard for them to give immediate and effective corrective feedback (Levis, 2007). Without targeted corrective feedback, simple mimicking and repetition do not work that well in class (Xu, 2017). However, the first few weeks of exposure to the Chinese language is critical in L2 Chinese learners' ability to produce correct tones partly because adult production and perceptual systems demonstrate only a certain degree of plasticity (Wang et al. 1999; Wang et al., 2003). L2 Chinese learners need to repeatedly practice listening to and producing tones since "good pronunciation of a language is a matter of motor skills, coupled with ear training" (Hockett, 1951). Therefore, it would be beneficial to find a tool that provides many opportunities for practice, that caters to each learner's need, and that is attractive enough that learners will actually use it.

Computer-assisted language learning (CALL) might meet these requirements. It has been an effective pedagogical approach since it was first integrated into American foreign language pronunciation teaching in the 1980s. CALL offers many advantages to the modern foreign language classes, including increasing learning efficiency and effectiveness, and providing easy access, great convenience, strong motivation, and institutional efficiency (Hubbard, 2009).

Short-term training using CALL, or CAPT (computer-assisted pronunciation teaching), has proven to be effective in assisting learners in the acquisition of new phonetic contrasts that do not exist in the phonological systems of their native language, be it English, Chinese, German, Cantonese, Spanish, or any other languages (Logan, Lively, & Pisoni, 1991; Wang et al., 1999; Wang et al., 2003; Kingston, 2003; Francis, Ciocca, Ma, & Fenn, 2008; Herd, Jongman, & Sereno, 2013; Li, 2016; Godfroid, Lin, & Ryu, 2017; Xu, 2017). In their studies, Wang et al. (1999) and Wang et al. (2003) investigated the effect of short-term auditory training on tonal performance in monosyllabic Chinese words while Li (2016) examined improvement in L2 learners' tonal perception in both mono- and disyllabic words after two weeks of perceptual training. After only eight sessions of auditory training, Wang et al. (1999; 2003) found that English-speaking learners of Chinese had improved significantly both in their tonal perception (by 21%) and tonal production (by 18%) on monosyllabic words. Li (2016), using the same test stimuli as in Wang et al. (1999, 2003) conducted a two-week auditory training on L2 learners' tonal perception of both

monosyllabic and disyllabic words. She found an overall 12% significant improvement after the auditory training. Moreover, learners showed greater improvement on disyllabic stimuli (12%) over monosyllabic stimuli (8%). All these studies indicated a significant gain in Chinese language learners' perception of both monosyllabic and disyllabic tones. Godfroid, Lin, and Ryu (2017) compared the effectiveness of using three single-cue methods and two dual-cue methods in helping L1 English L2 Chinese learners improve their perception of Chinese tones. Participants improved an average of 16% and 17% on trained syllables and 12% and 10% on untrained syllables. Only monosyllables (CV or CVC-structured) were used in their training and on tests.

However, those training sessions were all conducted in a clinical-lab environment that required an extra time commitment from L2 learners, facility support from the institute, and more management from the language instructor. Many scholars have recommended more research in natural classroom settings (Chun, 2015; Chun, 2017; Derwing & Munro, 2015). Chun (2017) pointed out that, although experimental studies are still needed to examine the effectiveness of application, it is as important to conduct studies "in authentic learning situations, for example, as part of regular scheduled language courses." But such studies are scarce.

Previous studies that investigated the acquisition order of the four tones by L2 Chinese learners have reached, more or less, a consensus that L2 learners acquire the level tones (e.g. T1) earlier than the contour tones (T2, T3, and T4). As for the contour tones, learners acquired the falling contour tone (T4) earlier than the rising contour tone (T2), which is, in turn, acquired earlier than the dipping tone (T3) (Zhang, 2018). Do training effects on tone perception vary among individual tones? Previous studies have shown mixed results. Wang et al. (1999) found no significant difference among the four tones on either the pre-test or the post-test, although, numerically, trainees' tones improved in the following order T4 > T2 > T3 > T1. So (2006) found that trainees needed more training sessions to distinguish between T1/T4 and T2/T4 than they did on T1/T2 or T1/T3. Would similar patterns of training effects be observed if the training is self-paced in a natural classroom setting? Would the online application facilitate L2 Chinese learners' acquisition of the harder to acquired tones, i.e. the contour tones?

Xu (2017) incorporated an online tone training application available at http://www.pinyinpractice.com into her beginning Chinese class. After two 20-minute sessions of in-class practice, her L2 Chinese learners' tonal perception improved significantly, as shown on the immediately administered post-test. Xu also took a survey of the learners' user experience of this online application and found that it was well received by a majority of the students. However, because Xu did not include a control group for comparison, it remains unclear whether the significant improvement was due to natural learning or was the benefit of the online training. Moreover, Xu only tested the perception of tones in monosyllabic words, which leaves the question of whether such training might be as helpful to students in perceiving tones in disyllabic words unanswered. For these reasons, it was necessary to re-investigate the effectiveness of online tone training practice through a new study that included a comparison group and tested the training's effectiveness on both monosyllabic words and disyllabic words. The acquisition order of

tones, training effects on different tones and learners' attitude towards such online tone training session was investigated as well.

## 3. The Current Empirical Study

### 3.1 Purpose of This Study

The present study explored the effectiveness of computer-assisted online tone practice in improving beginning L2 Chinese learners' perception of tones on monosyllabic and disyllabic words. It also explored whether the effect varied among different tones. The goal of this study was to identify an effective alternative, one that Chinese language instructors could easily incorporate into the curriculum. A group of students who received the traditional instructor-led instruction was used as a baseline for comparison. The following research questions were addressed:

1. Do L2 Chinese learners who receive online tone training in class show equal acuity in tone perception compared to students who receive traditional instructor-led tone training in class?
2. Do L2 Chinese learners who receive online tone training in class improve differently in tone perception on monosyllabic words or on disyllabic words than those who receive traditional instructor-led tone training in class?
3. Do L2 Chinese learners who receive online training in class improve differently on their perception of different tones on monosyllabic words or on disyllabic words than those who receive traditional instructor-led tone training in class?
4. How do students feel about online tone training?

Questions 1 through 3 were investigated by comparing the performance of two groups of students, one group receiving traditional instruction and the other using online tone-training practice, on the three tone perception tests (i.e. pre-test, post-test, and delayed post-test). A delayed post-test was used in order to measure the long-term effects that online tone training practice may have on Chinese learners' tonal perception. Neither the experimental group nor the control group was aware that they would be given a delayed post-test after they completed the post-test. Question number 4 was answered by analyzing answers to the questionnaire given to the students in the experimental group after they completed the delayed post-test.

The hypotheses regarding each research question that require statistical tests are listed as follows:

1. Research Question 1. Null hypothesis: there is no difference between the two groups in their total acuity of tone perception across tests. Alternative hypothesis: there is a difference between the two groups in their total acuity of tone perception across tests.
2. Research Question 2. Null hypothesis: there is no difference between the two groups in the acuity of their tone perception in the monosyllabic or the disyllabic settings across tests. Alternative hypothesis: there is a difference between the two

groups in their acuity of their tone perception in the monosyllabic or the disyllabic settings across tests.

3. Research Question 3. Null hypothesis: there is no difference between the two groups in the acuity of their perception of each tone in the monosyllabic or the disyllabic settings across tests. Alternative hypothesis: there is a difference between the two groups in the acuity of their perception of each tone in the monosyllabic or the disyllabic settings across tests.

## 3.2 Method and Design

### 3.2.1 Participants

Participants in this study were students enrolled in elementary Chinese classes at two universities in the American Midwest. The textbooks used by the two programs were *Integrated Chinese* and *Chinese Link* respectively. Despite the different textbooks used by the two programs, the grammar and vocabulary for first-year Chinese language learning were not much different and, in fact, are not much different across many textbooks (Cai & Sciban, 2008). We used natural class sessions and randomly decided on the division of the participants into the experimental group and the traditional group. Initially, 49 students participated in the study, but 14 of those students were ultimately excluded from the analysis due to previous Chinese language learning experience or exposure to other tonal languages. In addition, 4 more students were excluded from the analysis because they missed either the pre-test or the post-test. As a result, the data from 31 native English-speaking learners of Chinese were analyzed in this study, with 20 in the experimental group and 11 in the traditional group. These participants represented a variety of majors including business, finance, and political science. None had any history of hearing, speech, or language difficulties.

### 3.2.2 Training Materials and Test Instruments

All participants received a language learning background information sheet (see Appendix A) on the first day of class. The purpose of this information sheet was to ascertain whether participants had studied Chinese before or had had previous exposure to Chinese or any other tonal language, which might influence their performance on the tone perception task.

We chose http://www.pinyinpractice.com/ as the online tone training tool for the experimental group based on the selection criteria for CAPT tools proposed in Levis (2007). According to Levis (2007), a CAPT tool should provide a substantial amount of input, immediate and useful feedback, and motivate the learners to practice. The website www.pinyinpractice.com provides ample randomized tone exercises accompanied by audio files. Users can either choose the tones from the screen or enter the tones using numbers. It also provides instant feedback by indicating "correct" or "incorrect" in Chinese characters when learners pick or supply a tone for the syllable they hear in the audio. The interface is so well designed that learners are motivated to use it on their phone at any time that is convenient. Students in Xu's (2017) study expressed favorable feelings about the website. Moreover, the website is free and thus adds no financial burden to the students or

the instructors. This makes it more likely to be adopted in Chinese language courses if proven to be effective.

To measure the effectiveness of this online application, a perception task with 64 monosyllabic stimuli and 32 disyllabic stimuli adapted from Li (2016) was designed. The monosyllabic stimuli included all possible combinations of various initial consonants and final vowels, and different syllabic structures in Mandarin Chinese (i.e. V, CV, CVNasal, VN, CGlideV, and CGVN). In contrast, each disyllabic stimulus was composed of a combination of two random individual syllables from the monosyllabic stimuli. In such a design, the identical tones and vowel consonant combinations in both monosyllabic and disyllabic words ensured that students' perception of tones in both settings would be comparable. All stimuli were produced and recorded in a language lab by a well-trained female Chinese language instructor who had not been apprised of the purpose of the study and who had no contact with any of the students in either of the two Chinese programs. Pre-test, immediate post-test and delayed post-test were identical across both programs. The same sets of audio files were randomized on the pre-test, immediately administered post-test, and delayed post-test. For each version of the test, there was a corresponding test sheet composed of the syllables without tones (See Appendix B-1 for the test stimuli). For each test, students received a test sheet (Appendix B-2[1]) and were asked to circle the correct tone(s) for each stimulus according to what they heard on the audio recording. In each setting, monosyllabic or disyllabic, there are four tone marks provided after each syllable. Therefore, participants had to focus on the tones in order to choose the correct tone mark.

A survey was designed and conducted along with the delayed post-test with the students in the experimental group to investigate Chinese learners' attitudes toward tone learning and the use of online applications in learning tones. The survey consisted of self-assessment questions targeting the difficulties learners have in perceiving the tones, strategies that students believe can help improve their tone perception, and their experience with the use of the online application. The self-assessment questions were included because a good self-awareness of learning difficulties may help learners do more targeted practice if they are to use the online applications on their own in the future.

### 3.2.3 Procedure

All participants in both the experimental group and the traditional group completed a language learning background information sheet on the first day of Chinese class. A tone perception pre-test was given to the participants on the second day of class immediately following a short introduction to the four tones in Mandarin Chinese. The purpose of this test was to provide a baseline against which to compare participants' improvement. One group of participants received traditional tone practice while the other did online tone training practice in class which consisted of four 15-minute sessions spread over a two-week period. The online tone practice included both monosyllabic and disyllabic words

[1] Appendix B-2 shows the test sheet that is used for the pre-test. The test sheets used for the immediate post-test and the delayed post-test were different only in the order of test items, which are not included in the appendix to save space.

found on the website www.pinyinpractice.com, with the first two sessions focusing on monosyllable tone practice and the last two sessions on disyllable tone practice.

A tone perception post-test was administered one day after the last training session. Four weeks later, a delayed tone perception post-test was administered to both groups. Learners in the experimental group were also asked to complete a questionnaire that addressed their experience with and their attitudes towards the online application.

During the time that the experimental group practiced tones in class using www.pinyinpractice.com, the traditional group practiced tones in class with the instructors in the traditional way, namely, through repetition after the instructor and reading aloud, both followed by immediate corrective-feedback from the instructor. For instance, the instructor usually pronounced a syllable with the correct tones after which the students repeated the syllable in groups or individually. If a student made an error, whether it was on the initial, final or the tone, the instructor would correct the student by pronouncing the syllable correctly and asking him/her to repeat.

### 3.2.4 Data Coding

All data from the language learning background information sheets, the three tone perception tests, and the tone learning experience questionnaire were input into Excel files. Data were cross-checked by two of the authors to ensure accuracy.

The points for the 64 monosyllabic stimuli and the 32 disyllabic stimuli were collected. One point was given for each correct tone in each section of the tone perception tests. One token from the monosyllable section and two from the disyllable section were removed from the final analysis due to some ambiguity. As a result, the total number of points available from the monosyllable section was 63, and the total number from the disyllable section was 60. The total number of points received by each participant from each test was converted into an accuracy rate to make the results comparable across different sections and different tests.

The coding of the semi-open-ended questionnaire was done using a mixed method. For those questions with multiple choice answers, the percentage of each choice was calculated by compiling all the answers to each question and then calculating the percentage of each choice. For the open-ended questions, all answers were compiled and read through to identify patterns and categorize them.

## 3.3 Results

In this section, the results of the tone perception tasks from the pre-test, immediately administered post-test, and delayed post-test and the results from the questionnaire will be reported.

### 3.3.1 Results of Tone Perception Tests

Tone perception accuracy rates were used as the measurement on all tone perception tests. An overall score from both the monosyllabic words section and disyllabic words section was calculated for each participant.

#### 3.3.1.1 Overall Results of Tone Perception Tests

The overall accuracy rates of both groups on the pre-test, immediately administered post-test, and delayed post-test are shown in Figure 2 and Table 1. These results answer the first research question. Both groups showed improvement in their perception of tones from pre-test to the immediately administered post-tests. The tone perception accuracy rate of the experimental group increased from 60.9% on the pre-test to 78.8% on the immediately administered post-test and 85.9% on the delayed post-test. Similarly, the tone perception accuracy rate of the traditional group increased from 53.8% to 76% on the immediately administered post-test and 78.6% on the delayed post-test. The experimental group performed better on perceiving tones than did the students in the traditional group in absolute value. In order to see whether the differences between the experimental group and the traditional group within each test were statistically significant, a one-way repeated measures ANOVA was conducted with three levels: pre-test, immediately administered post-test, and delayed post-test. The results of homogeneity of variances tests were not significant at any level, which indicates that the two groups were comparable at any level. However, the null hypothesis of sphericity in Mauchly's test was rejected (p=.006), therefore, the Greenhouse-Geisser adjustment was used. The results indicated that there was no significant interaction between the time and group: $F(1.53, 44.34) = .872$, $p = .399$. Therefore, the null hypothesis was accepted for research question 1. In other words, there was no difference between the two groups in their total acuity of tone perception across tests.

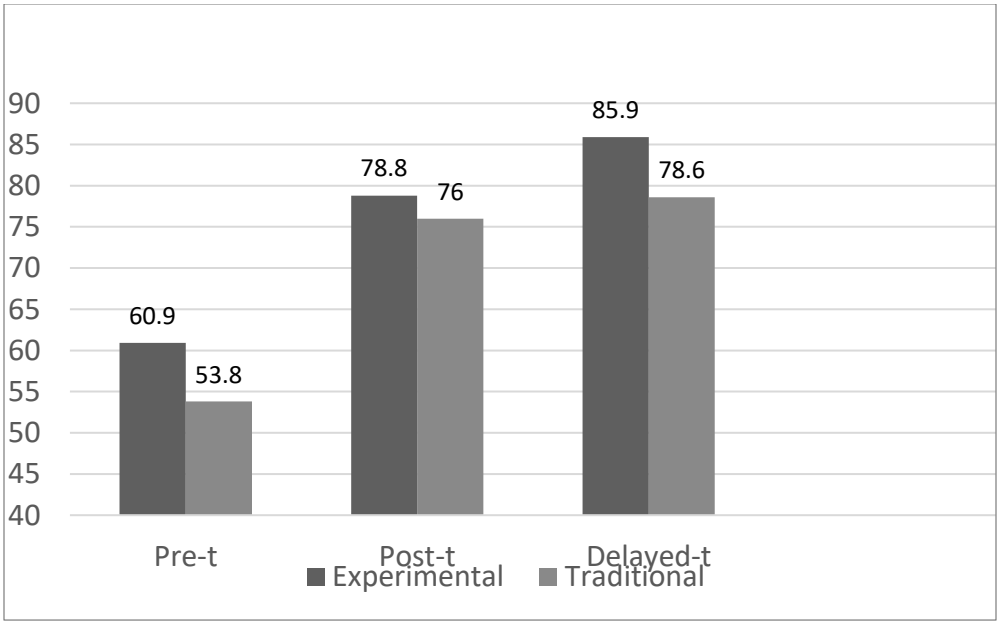



**Figure 2 Overall Tone Perception Accuracy Across Groups**

**Table 1 Group Perception Accuracy Across Time: Total Scores**

| Time | Group | Pre-test | | Post-test | | Delayed-post-test | |
|---|---|---|---|---|---|---|---|
| | | M | SD | M | SD | M | SD |
| **Total** | Experimental | 60.9 | 13.0 | 78.8 | 9.2 | 85.9 | 5.8 |
| | Traditional | 53.8 | 14.1 | 76.0 | 9.6 | 78.6 | 7.4 |

### 3.3.1.2. Tone Perception on Monosyllabic Words and Disyllabic Words

The second research question focused on participants' perception accuracy in two different settings. Figures 3 and 4 show the tone perception accuracy rates on monosyllabic words and disyllabic words respectively, and detailed descriptive statistics are summarized in Table 2. As can be seen in Figure 3, both groups of learners performed at an accuracy rate of over 68.3%, even on the pre-test on monosyllabic words after only a brief introduction to tones. The accuracy rate increased to over 92.4% on the immediately administered post-test and to over 94.4% on the delayed post-test.

Figure 4 shows participants' accuracy rates on tones of disyllabic words. Compared to learners' performance in a monosyllabic setting, their performance in the disyllabic setting was not ideal. On the pre-test, the accuracy rate was 38.6% for the traditional group, and 45.8% for the experimental group. On the immediately administered post-test, the accuracy rate for the traditional group was 58.8% and 63.5 % for the experimental group, and the accuracy rate increased only slightly on the delayed post-test (62.1% for traditional group and 73.2% for the experimental group).

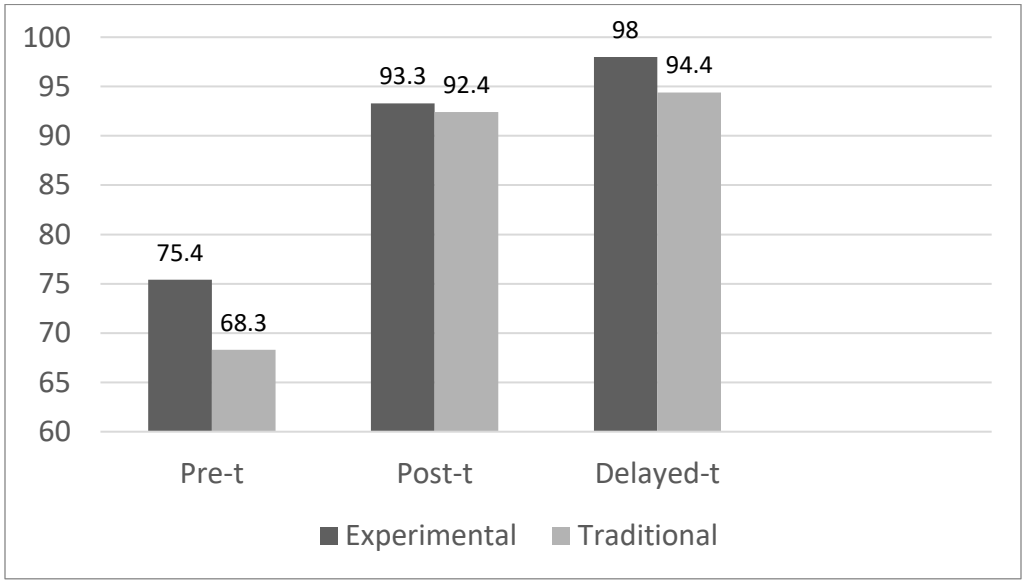



**Figure 3 Perception Accuracy on Monosyllabic Words Across Groups**

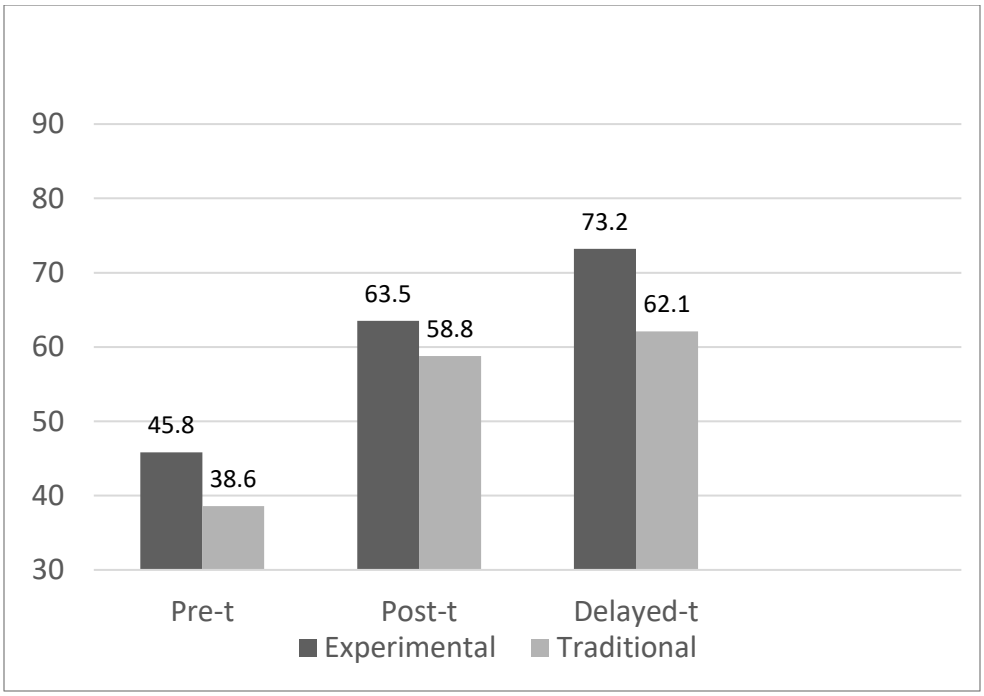



**Figure 4 Perception Accuracy on Disyllabic Words Across Groups**

**Table 2 Group Perception Accuracy Across Time in Different Settings**

| Time | Group | Pre-test | | Post-test | | Delayed-post-test | |
|---|---|---|---|---|---|---|---|
| | | M | SD | M | SD | M | SD |
| **Monosyllabic words** | Experimental | 75.4 | 13.6 | 93.3 | 8.3 | 98.0 | 2.2 |
| | Traditional | 68.3 | 19.6 | 92.3 | 6.7 | 94.4 | 6.0 |
| **Disyllabic words** | Experimental | 45.8 | 14.3 | 63.5 | 12.2 | 73.2 | 11.2 |
| | Traditional | 38.6 | 9.6 | 58.8 | 13.9 | 62.1 | 9.5 |

In order to test if there were any significant differences between the two groups' performances on monosyllables and disyllables across tests, a doubly multivariate repeated measures ANOVA was conducted. There were repeated measures over time (pre-test, post-test, and delayed post-test), the independent variable was the group, and the dependent variables were their accuracy scores for monosyllabic and disyllabic words. The results indicated that overall there was no significant difference between the two groups in the acuity of their tone perception in the monosyllabic or disyllabic settings across tests: Wilks' Lambha = .935, $F(4, 114) = .977$, $p = .423$. When broken down into different settings, there was no significant difference between the groups either. In the monosyllabic setting, Mulchly's test rejected sphericity ($p < .001$), so the Greenhouse-Geisser-adjusted results were used. There was no interaction effect between the time and group: $F(1.33, 38.65) = .654$, $p = .466$. In the disyllabic setting, Mulchly's test of sphericity was not significant ($p = .186$). There was no interaction effect between the time and group on tones for disyllabic words: $F(2, 58) = 1.248$, $p = .295$. Therefore, the second null hypothesis was accepted. In other words, the results indicated that both groups showed a similar pattern in their performance in the two different settings across the tests.

### 3.3.1.3. Tone Perception Accuracy on Different Tones

The results of research question 3 are reported in this section. To discover whether online tone training has an equal effect as the traditional training on different tones, participants' tone perception accuracy in the monosyllabic setting and disyllabic setting was further broken down by tone. Table 3 summarizes the groups' tone perception accuracy on different tones across the tests in the monosyllabic setting. On the pre-test, the experimental group performed best on perceiving the 1st tone (93.8%) and the traditional group performed best on perceiving the 3rd tone (79%). On the immediately administered post-test, both groups performed best on perceiving the 1st tone: the experimental group achieving a 98.1% accuracy rate and the traditional group 94.9%. Both groups recorded their worst performance in perceiving the 2nd tone with the experimental group achieving an 85.7% accuracy rate and the traditional group 85.5%. The same pattern was displayed on the delayed post-test: both groups performed the best in perceiving the 1st tone (the experimental group 99.7%; the traditional group 98.3%) and the worst in perceiving the 2nd tone (the experimental group 96.3%; the traditional group 87.9%).

**Table 3 Group Perception Accuracy on Different Tones in Monosyllabic Words across Tests**

| Time | Group | Pre-test | | Post-test | | Delayed-post-test | |
|---|---|---|---|---|---|---|---|
| | | M | SD | M | SD | M | SD |
| **Tone 1** | Experimental | 93.8 | 8.1 | 98.1 | 7.1 | 99.7 | 1.4 |
| | Traditional | 73.9 | 25.0 | 94.9 | 6.1 | 98.3 | 4.0 |
| **Tone 2** | Experimental | 51.3 | 26.5 | 85.7 | 20.2 | 96.3 | 5.1 |
| | Traditional | 49.7 | 19.6 | 85.5 | 12.6 | 87.9 | 19.0 |
| **Tone 3** | Experimental | 84.7 | 17.4 | 95.3 | 7.6 | 97.5 | 4.3 |
| | Traditional | 79.0 | 20.4 | 93.2 | 8.6 | 93.8 | 8.4 |
| **Tone 4** | Experimental | 70.3 | 25.6 | 94.4 | 8.3 | 98.4 | 4.9 |
| | Traditional | 68.2 | 28.2 | 92.0 | 10.5 | 97.2 | 4.3 |

Participants' average accuracy in perception of different tones in the disyllabic setting is shown in Table 4. On the pre-test, both groups performed the best on perceiving the 1st tone with the experimental group achieving an accuracy rate of 55% and the traditional group 42.6%. Both groups performed the worst on the 2nd tone with the experimental group achieving an accuracy rate of 38.8% and the traditional group 31.8%. On the immediately administered post-test, both groups still performed the worst on the 2nd tone (experimental group 52.5%; traditional group 51.7%), and the experimental group still performed the best on perceiving the 1st tone (73.8%), while the traditional group performed the best on the 4th tone (65.9%). On the delayed post-test, both groups performed the best on the 4th tone (the experimental group 81.9%; the traditional group 69.9%). The experimental group performed the worst on the 3rd tone (60.4%) while the traditional group performed the worst on 2nd tone (51.7%).

**Table 4 Group Perception Accuracy on Different Tones in Disyllabic Words Across Tests**

| Time | Group | Pre-test | | Post-test | | Delayed-post-test | |
|---|---|---|---|---|---|---|---|
| | | M | SD | M | SD | M | SD |
| **Tone 1** | Experimental | 55.0 | 18.1 | 73.8 | 15.5 | 78.4 | 18.4 |
| | Traditional | 42.6 | 12.4 | 64.2 | 19.8 | 67.0 | 14.0 |
| **Tone 2** | Experimental | 38.8 | 17.2 | 52.5 | 20.6 | 68.4 | 13.1 |
| | Traditional | 31.8 | 8.6 | 51.7 | 24.5 | 51.7 | 22.0 |
| **Tone 3** | Experimental | 45.4 | 22.0 | 54.6 | 17.2 | 60.4 | 16.0 |
| | Traditional | 34.8 | 18.2 | 52.3 | 18.3 | 53.8 | 9.4 |
| **Tone 4** | Experimental | 43.8 | 18.8 | 71.6 | 16.7 | 81.9 | 15.3 |
| | Traditional | 43.8 | 21.8 | 65.9 | 21.7 | 69.9 | 20.7 |

In order to test if the two groups performed differently among the tones in the monosyllabic and the disyllabic settings, a doubly multivariate repeated measures ANOVA was conducted. There were repeated measures for the tones (T1, T2, T3, and T4) and the time (pre-test, post-test, and delayed post-test), the independent variable was the group, and the dependent variable was their accuracy scores for both settings (monosyllabic vs. disyllabic).

The multivariate test results indicated that overall there was no significant difference between the two groups in the acuity of their perception of different tones in the monosyllabic and disyllabic settings across tests, because there was no significant interaction effect between tone, time, and group: Wilks' Lambha = .910, $F(12, 346) = 1.394$, $p = .166$. The reason being that when broken down into different settings, there was no significant interaction between tone, time, and groups for the univariate tests. In the monosyllabic setting, Mulchly's test of sphericity was significant ($p < .001$), therefore the Greenhouse-Geisser adjustment was used. There was no interaction effect among tone, time, and group: $F(3.63, 105.21) = 1.905$, $p = .383$. In the disyllabic setting, Mulchly's test of sphericity was not significant ($p = .538$) and there was no interaction effect between tone, time, and group either: $F(6, 174) = 1.069$, $p = .383$. Therefore, the third null hypothesis was accepted and the alternative hypothesis was rejected. In summary, the results indicated that both groups showed a similar pattern in their performance on different tones in the two different settings across the tests.

Although no significant interaction effect was found between the groups across the tones, an interesting trend was noticed regarding T2 in both monosyllabic and disyllabic settings. After we singled out the accuracy rates of T2 and compared the change of each group's acuity across the tests, the following graph (Figure 5) was created:

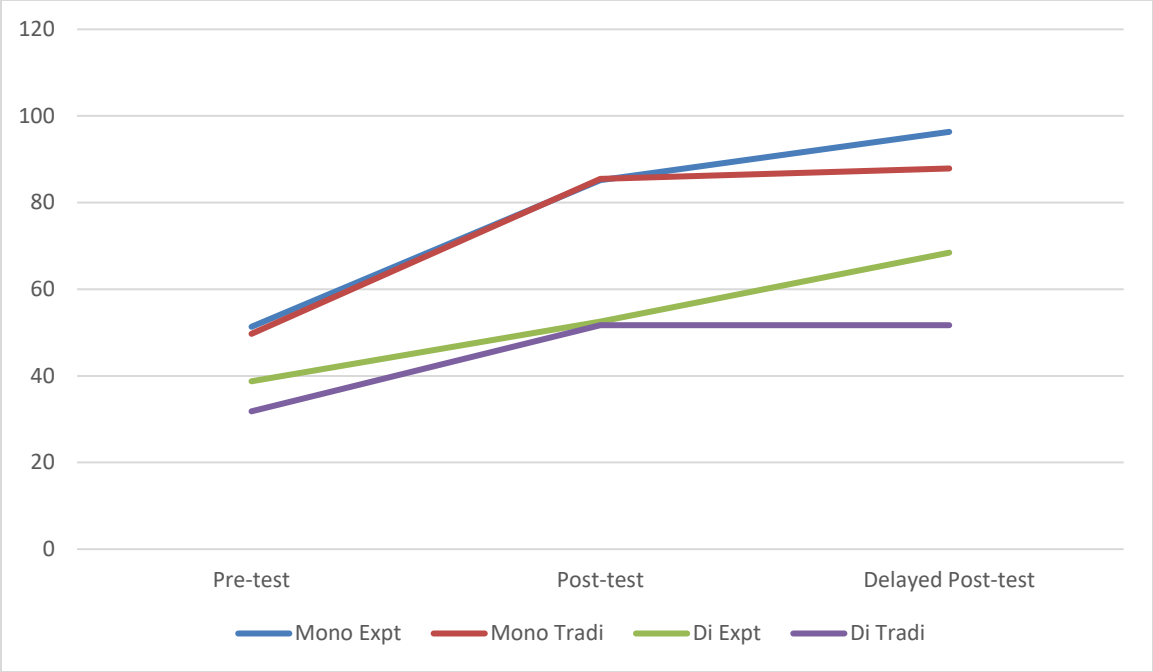



**Figure 5 Perception Accuracy of T2 in Different Settings Across Time**

Both groups started around the same accuracy, and they reached similar accuracy at the end of the training sessions. However, the experimental group's accuracy kept improving, whereas the traditional group's accuracy stopped increasing.

### 3.3.2 Questionnaire Results

The questionnaire focused on Chinese learners' self-assessment of the difficulties they have in perceiving tones, their tone learning strategies, their experience with the online tone training practice, and their preference regarding using online or traditional tone training practices in the classroom. Only learners in the experimental group completed the questionnaire. Because the questionnaire is qualitative in nature, 24 responses[2] were included in the final analysis.

#### 3.3.2.1 Students' Perception of Different Tones

Part I of the questionnaire elicited students' perception of different tones: 62.5% of the participants rated the 1st tone as the easiest one to identify while 25% of the participants rated the 3rd tone and 12.5% rated the 4th tone as the easiest. Nobody rated the 2nd tone as easy. As for the most difficult tone to identify: 58% of the participants rated the 2nd tone and 25% of participants rated the neutral tone as being most difficult to accurately perceive while 8.3% of participants rated the 3rd and 4th tones as being the hardest to identify. This

[2] Four of these subjects' tone perception data were not included in the analysis of tone perception tests because they either missed the pre-test or the post-test. We included their questionnaire responses here because we believe the questionnaire is qualitative in nature and these subjects' experiences using the online tone training websites are as valuable as the other Chinese learners in the experimental group.

was consistent with students' performance on the tone perception tasks: their accuracy was highest on the 1st tone and lowest on the 2nd tone. These results are in line with the results of other experimental studies that found the 2nd tone and the 3rd tone to be harder than the 1st or 4th tone for students to perceive after training (Wang et al., 1999; Wang et al., 2003; Li, 2016; Sun, 1998; He & Wayland, 2013). Our participants also mentioned that the neutral tone was difficult to perceive. However, the neutral tone was not included in our study.

#### 3.3.2.2 Student Strategies for Learning Tones

Part II of the questionnaire collected the strategies employed by students for learning tones. Not surprisingly, students employed a variety of strategies. Table 5 shows the percentage of students who used each strategy. Since students could check multiple items, the total number of strategies was tallied by adding up all items the participants had checked, then calculating the percentage for using a certain strategy.

The strategy used by most students (87%) was "listening to audio files" followed by "using online resources to practice" (83.3%). Almost 80% of the participants used "comparing what I hear with what I say" and 50% of the participants used the strategy "repeating after audio files." In addition, 29.2% of the participants employed a strategy of "practicing listening and reading pinyin with fellow students." Strategies other than those listed in the questionnaire were employed by 20.8% of participants.

**Table 5 Student Strategies for Learning Tones**

| Tone learning strategies | Frequency percentage |
|---|---|
| Listening to audio files | 87.5% |
| Using online resources to practice | 83.3% |
| Comparing what I hear with what I say | 79.2% |
| Repeating after audio files | 50% |
| Practicing listening to and reading pinyin with fellow students | 29.2% |
| Other | 20.8% |

#### 3.3.2.3 Students' Computer-assisted Language Learning (CALL) Experience

Part III of the questionnaire focused on students' computer-assisted language learning (CALL) experience. All participants liked the website www.pinyinpractice.com, and thought it was both useful and helpful, especially when it came to improving their perception of tones in disyllabic words where tones are not as obvious in actual speech as they are when demonstrated by an instructor whose enunciation is more deliberate. Some participants felt frustrated using the website, especially at the beginning, but with greater familiarity came greater confidence. A large majority (83.3%) of the participants reported that they used the website outside of class for 15-30 minutes, two to three times a week.

Less than half (41.6%) of the participants also tried other features offered by www.pinyinpractice.com, although they had not been instructed to do so. A little over half

(54.2%) of the participants expressed an interest in trying other features in the future. However, 58.3% of the participants indicated that they would prefer to practice other features in class before using them outside of class.

In general, students' attitudes towards www.pinyinpractice.com were positive. They liked the set-up and the simple straightforward interface. They also appreciated other features such as instant feedback, the syllable replay option, and the variety of input methods (choosing or entering tones). At the same time, the participants also offered their thoughts on improvements that they would like to see on the website. These included an option for revealing the correct tones after a few attempts. The current interface only gives very general feedback such as "right" or "wrong," and does not reveal the correct answer until the learner gets it right. Nor does it give feedback on any specific error that students might have made. This vagueness sometimes caused anxiety and frustration for the students. Other suggestions included a way for the missed tones to be marked in the instant feedback, the ability to adjust the speed at which each syllable is read, and an option to enter tones in a way other than using numbers, such as to input the tone marks as diacritics.

When participants were asked whether they preferred traditional in-class instructor-led tone practice or self-paced on-line tone practice, 67.57% of the participants said they preferred traditional in-class tone practice while only 32.43% preferred online practice. Students who preferred in-class instructor-led tone practice appreciated the interactive nature of this method, pointing out that such practice allowed them to say the syllable aloud and get instant feedback from the instructors. Students who preferred the online tone practice liked the fact that they could practice at their own pace, and practice as many times as they wanted to without feeling intimidated by the perceived judgment of their instructor or classmates when they made mistakes.

### 3.4 Discussion

The first research question examined the overall performance of participants who received online training as compared to those who received traditional in-class tone training. The results of the tone perception tests indicate that L2 Chinese learners who received the online tone training in class showed equal acuity in tone perception compared to students who receive traditional instructor-led tone training in class. In terms of gains in acuity, the experimental group showed an increase of 18.1% on the immediate post-test and 25% on the delayed post-test respectively. This finding is similar to the training gains reported in Godfroid et al. (2017) of 10-20%. All participants in the experimental group and the control group practiced tones during the training period. The difference between the traditional group and the experimental group lay in the format of such tone training, i.e. instructor-led vs. self-paced online tone practice. Both groups showed improved tone perception after the training period was over as can be seen on the immediate post-test. Therefore, we can see that tone training, no matter what the format might be, has an immediate effect on tone perception, a finding which is consistent with those reported in tone training literature (Wang et al., 1999, 2003; Li, 2016; Godfroid, et al. 2017). The online self-paced tone practice helped the students as much as the instructor-led tone practice as far as their tone perception is concerned.

The second research question examined whether L2 Chinese learners who received the online tone training in class improve differently in tone perception on monosyllabic words or on disyllabic words than those who received traditional instructor-led tone training in class. The results indicated that the experimental group's acuity improved significantly over time in both settings and they showed a similar pattern of improvement as the traditional group. The two groups did not differ significantly from each other in the acuity of their tone perception in the monosyllabic or the disyllabic settings across tests. The results also indicated that both groups' tone perception accuracy on monosyllabic tokens was higher than their accuracy on disyllabic tokens. This finding is consistent with findings from previous studies, i.e. the tones of monosyllabic words are easier for Chinese learners to identify than those of disyllabic words (He & Wayland, 2013; Hao, 2012; Li, 2016). According to Xu (1997), the tones of monosyllabic words are more canonical and stable while tones of disyllabic words are always affected by adjacent tones making them more difficult to discern. The two groups of learners showed the same patterns in their performance in perceiving the tones of monosyllabic words and disyllabic words. This indicates that tone-training practice, no matter whether it is online or in-class instructor-led, is effective in helping L2 Chinese learners correctly perceive tones in both monosyllabic and disyllabic words.

The third research question addressed whether students who received online tone training showed a similar pattern as their peers who received traditional instructor-led tone training in their perception across the tones in both the monosyllabic and disyllabic settings. The tone perception tests' results indicate that improvement was not just dependent on time but also on specific tone: students performed very well in identifying the T1 on the pre-test even after just a short introduction to tones, but they did not do as well in perceiving T2, T3, and T4. Through training, they improved their perceptions on T2, T3, and T4 in the monosyllabic setting greatly, as well as in the disyllabic setting. However, the pattern of improvement of different tones across time looked similar between the online application experimental group and the traditional instruction group. What is noteworthy is that for T2, the experimental group' acuity continued to improve after the training ended in both the monosyllabic and disyllabic settings, whereas the traditional group' acuity stopped improving, even though the difference is not significant. Since T2 was the most difficult tone to perceive as indicated from both students' performance in the tests and their self-evaluation in the questionnaire, the results are very encouraging.

Why did the online application experimental group keep improving their perception of T2? Based on the information provided by the questionnaire, we offer this possible explanation: while the traditional group stopped practicing tones after the classroom tone-training period was over, some participants in the experimental group continued using the website to practice tones. A majority of the participants reported that they used the website outside of class for 15-30 minutes, two to three times a week, although the participants did not indicate whether they devoted extra time to practice outside of class during the experiment period or after it. Regardless of when they practiced, the significance of introducing the online tone practice website to the students in the classroom not only has immediate, short-term, tone training benefits, but also has extended, long-term benefits, particularly as a tool students can use outside the classroom to help them develop autonomy. Although the reported extra use of the online practice made the present study less

controlled in the sense that the experimental group spent more time practicing their perception of tones, the fact that these participants did, in fact, practice outside class without being told to is encouraging. Another explanation could be that the self-paced online training gave students greater exposure to the tones that they found more difficult, making their training more targeted and individualized and thus increasing retention.

The fourth research question investigated students' experience using the online training website. Overall, the students had a positive attitude toward the pinyin practice website. They acknowledged that the website was helpful and a majority (83.3%) used online-training in class as well as outside the classroom. However, according to their responses, more students preferred in-class instructor-led tone practice to online self-paced practice (67.57% vs. 32.43%). Several factors may contribute to this preference. One of the most obvious differences between the traditional classroom instructor-led practice and the online tone-training practice is that students interact with human beings in the former while interacting with computers in the latter. Secondly, the feedback they get from the instructor is different than the feedback they get from the computer. Although both can be counted as immediate feedback, that from the instructor is more specific, targeting the errors that the students made in production, whereas the latter focuses only on whether the tone the student perceived was correct or not; in other words, the traditional classroom instructor-led practice focuses on production while the online tone-training is only about perception. Students may feel that they are more engaged in the traditional classroom instructor-led practice than in the online tone-training activities.

## 4. Conclusion and Pedagogical Implications

Overall, the current study shows that using online applications helped students perceive tones as effectively as the traditional instructor-led training in class in terms of overall acuity, acuity in monosyllabic or disyllabic words, or acuity with individual tones. The introduction of online practice tools such as the one used in the current study is beneficial in that it gives students access to a platform on which they can —and are motivated to—continue working on their tone perception even after the formal pinyin training period has ended. It thus can have long-term positive effects in the training of students in correct tone perception. The fact that students voluntarily used the website outside of the classroom to practice tones is also encouraging. Although the introduction of the tone-training website in class is an effective alternative, such online tone training practice is, perhaps, better integrated into the curriculum as an assignment rather than as a classroom activity. In other words, instructor-led practice in the classroom is still necessary and well-received, while the online applications can be used as an effective alternative to help students improve their tone perception outside of class since there is not enough time to be spent on tone training in class.

Although it has not been empirically proven, it can be inferred from the current study that assigning tone training three times a week for four weeks after the completion of the initial pronunciation practice could be beneficial. To consolidate the effects, a tone perception test given in the fifth week might be desirable.

## 5. Limitation and Future Studies

Although the authors did their best to recruit as many participants as possible, the sampling in this study was not large enough to reach more solid conclusions, especially when the large variation within each group was taken into consideration. The training tool used in this study was not perfect either. For example, it only gave a general feedback such as "right" or "wrong," without targeting any specific error that students might have made. This vagueness made the training less efficient and raised students' anxiety and frustration in some cases, as was indicated in the findings of the questionnaire. In the future, if a more efficient and targeted online application can be used to explore this issue, a better result may be obtained. In addition, recent studies on Chinese tone training show that perceptual training has a clear, facilitative effect on both the perception and production of tones (Leather, 1990; Wang et al., 1999; Wang et al., 2003). More studies on the tone perception training effects on tonal production are hence clearly called for.

## References

Cai, W., & Sciban, S.-N. (2008). How is Chinese grammar taught? A comparison of grammar materials in five Chinese textbooks. *Journal of the Chinese Language Teachers Association, (43)*1, 17-31.

Chun, D. M. (2015). Language and culture learning in higher education via telecollaboration. *Pedagogies: An International Journal, 10*(1), 5-21.

Chun, D. M. (2017). Research methods for investigation technology for language and culture learning. In C. Chapelle & S. Sauro (Eds.), *The handbook of technology in second language teaching and learning* (pp. 393-408). Hoboken, NJ: Wiley-Blackwell.

Chun, D. M., Jiang, Y., Meyr, J., & Yang, R. (2015). Acquisition of L2 Mandarin Chinese tones with learner-created tone visualizations. *Journal of Second Language Pronunciation*, *1*, 86-114.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research* (Vol. 42). Amsterdam: John Benjamins Publishing Company.

Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics*, *36*(2), 268-294.

Godfroid, A., Lin, C. H., & Ryu, C. (2017). Hearing and seeing tone through color: An efficacy study of web-based, multimodal Chinese tone perception training. *Language Learning*, *67*(4), 819-857.

Hao, Y-C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, *40*(2), 269-279.

He, Y., & Wayland, R. (2013). Identification of Mandarin co-articulated tones by inexperienced and experienced English learners of Mandarin. *Chinese as a Second Language Research*, *2*(1), 1-21.

Herd, W., Jongman, A., & Sereno, J. (2013). Perceptual and production training of intervocalic/d, ɾ, r/in American English learners of Spanish. *The Journal of the Acoustical Society of America*, *133*(6), 4247-4255.

Hockett, C. F. (1951). *Progressive exercises in Chinese pronunciation* (Vol. 2). Institute of Far Eastern Languages, New Haven: Yale University.

Hubbard, P. (Ed.). (2009). *Computer assisted language learning: Critical concepts in linguistics. Present trends and future directions in CALL*. New York: Routledge.

Jongman, A., Wang, Y., Moore, C. B., & Sereno, J. A. (2006). Perception and production of Mandarin Chinese tones. In P. Li, L. Tan, E. Bates, & O.J.L. Tzeng (Eds.), *The handbook of East Asian psycholinguistics: Volume 1, Chinese*, (pp. 209-217). Cambridge: Cambridge University Press.

Kingston, J. (2003). Learning foreign vowels. *Language and Speech, 46*(2-3), 295-348.

Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. *New sounds, 90*, 72-97.

Lee, C. Y., Tao, L., & Bond, Z. S. (2010). Identification of multi-speaker Mandarin tones in noise by native and non-native listeners. *Speech Communication, 52*(11), 900-910.

Levis, J. (2007). Computer technology in teaching and researching pronunciation. *Annual Review of Applied Linguistics, 27*, 184-202.

Li, Y. L. (2016). *Effects of high variability phonetic training on monosyllabic and disyllabic Mandarin Chinese tones for L2 Chinese learners*. Retrieved from ProQuest Dissertations and Theses Global. (1800269699).

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English/r/and/l: A first report. *The Journal of the Acoustical Society of America, 89*(2), 874-886.

Miracle, W. C. (1989). Tone production of American students of Chinese: A preliminary acoustic study. *Journal of the Chinese Language Teachers Association*, *24*(3), 49-65.

Orton, J. (2013). Developing Chinese oral skills: a research base for practice. In I. Kecskés (Ed.), *Research in Chinese as a second langua*ge (pp. 3-26). Berlin: De Gruyter Mouton.

Rochet, B. L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 379-410). Baltimore: York Press.

Rohr, J. (2014). Training naïve learners to identify Chinese tone: An inductive approach. In N. Jiang (Ed.), *Advances in Chinese as a Second Language: Acquisition and Processing*, (pp. 158-178). Newcastle: Cambridge Scholars Press.

Shen, X. S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of the Chinese Language Teachers Association, 24*(3), 27-47.

Shen, X. S., & Lin, M. C. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech, 34*, 145–156.

So, C. K. L. (2006). *Effects of L1 prosodic background and AV training on learning Mandarin tones by speakers of Cantonese, Japanese, and English*. Ph.D. dissertation, Simon Fraser University.

Sun, S. H. (1998). *The development of a lexical tone phonology in American adult learners of standard Mandarin Chinese* (No. 16). Honolulu: University of Hawaii Press.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *The Journal of the Acoustical Society of America*, *106*(6), 3649-3658.

Wang, Y., Jongman, A., & Sereno, J. (2003). Acoustic and perceptual evaluation of Mandarin tone production before and after training. *Journal of the Acoustical Society of America*, *113*, 1033-1043.

Xing, J. Z. (2006). *Teaching and learning Chinese as a foreign language: A pedagogical grammar*. Hong Kong: Hong Kong University Press.

Xu, H. (2017). *Self-paced use of online applications on Chinese tone perception among beginners in Chinese*. Paper presented at the 2017 UW Spring Conference on Teaching and Learning, La Crosse, WI.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, *25*(1), 61-83.

Zhang, H. (2018). Current trends in research of Chinese sound acquisition. In C. Ke (Ed.), *The Routledge handbook of Chinese second language acquisition*. London & New York: Routledge Taylor & Francis Group.

**Appendix A: Language background information sheet**

English Name: _____ Email: _____

Phone#: _____ School year: _____

Major: _____

Please answer the following questions.

1. Do you have any experience studying Chinese? If the answer is yes, please specify where and when. *Please remember to talk to your lecture teacher if you have any background speaking Chinese, whether or not you have taken placement test.

2. Have you ever been to China? If the answer is yes, please specify where and when.

3. Do you speak any Chinese dialects such as Cantonese etc.?

4. Have you learned any other foreign languages (excluding Mandarin Chinese)? If have, how long?

5. Do you have a Chinese name? If you do, what is it? Please write it down in Chinese characters or describe the meaning of the characters. If you do not have a Chinese name, do you have any preference?

**Appendix B-1: Test stimuli used in the pre-test, post-test, and delayed post-test (the syllables were randomized in each test)**

*Monosyllables:*

| | | | |
|---|---|---|---|
| 1. ji ā | 2.ji á | 3.ji ǎ | 4.ji à |
| 5. w ān | 6.w án | 7.w ǎn | 8.w àn |
| 9. li ū | 10.li ú | 11.li ǔ | 12.li ù |
| 13.sh ēn | 14.shén | 15.shěn | 16.shèn |
| 17. xu ē | 18. xué | 19.xuě | 20.xuè |
| 21.tu ō | 22.tuó | 23. tuǒ | 24. tu ò |
| 25.ku ī | 26.tu í | 27.tuǐ | 28.tuì |
| 29.t āo | 30.táo | 31. tǎo 3 | 2.t ào |
| 33. huān | 34.huán | 35. hu ǎ n | 36. huàn |
| 37.shī | 38.shí | 39. shǐ | 40.shì |
| 41.wā | 42.wá | 43.wǎ | 44.w à |
| 45.xī | 46.xí | 47.xǐ | 48. xì |
| 49.chōng | 50.chó ng | 51.chǒng. | 52chòng |
| 53.yōu | 54.yóu | 55.yǒu | 56.yòu |
| 57. pū | 58.pú | 59.pǔ | 60.pù |
| 61. yāng | 62. yáng | 63. yǎng | 64. yàng |

*Disyllables*

| | | | |
|---|---|---|---|
| 1. xīshǐ | 2. huánliù | 3.pùxuē | 4.jiàkuǐ |
| 5. chóngshí | 6. tǎojiā | 7. xuěyòu | 8. tuōxí |
| 9. wǎtuǒ | 10.yóuliǔ | 11.tuóxǐ | 12.shěnhuān |
| 13. liūhuǎn | 14.púshī | 15.wāhuàn | 16.chǒngshì |
| 17.yǒushén | 18.jiǎwǎn | 19.wānyōu | 20.xìxué |
| 21.yángxuè | 22.wànwán | 23.jiákuí | 24.shèntuò |

25.yāngwà 26.pūchōng 27.yǎngwá 28.kuìshēn

29. tàopǔ 30.kuīliú 31.chòngyàng 32.táotāo

**Appendix B-2: Test sheet used in the study**

**Tone Perception**

Name: _____ Lecture time (circle one): 9AM/10AM

You are going to hear some syllables. Please indicate the tone of each syllable you hear by circling the tone marker provided. For example, if you see the following syllable: ma ( – / V \ ) , and then you hear: mă. You circle the third tone: "v": ma ( – / V \ ) .

**Part I: Monosyllables**

1. jia ( – / V \ )
2. shi ( – / V \ )
3. shen ( – / V \ )
4. kui ( – / V \ )
5. xue ( – / V \ )
6. shen ( – / V \ )
7. huan ( – / V \ )
8. xi ( – / V \ )
9. xue ( – / V \ )
10. shen ( – / V \ )
11. wa ( – / V \ )
12. wan ( – / V \ )
13. yang ( – / V \ )
14. shi ( – / V \ )
15. chong ( – / V \ )
16. yang ( – / V \ )
17. liu ( – / V \ )
18. shi ( – / V \ )
19. xue ( – / V \ )
20. huan ( – / V \ )
21. jia ( – / V \ )
22. huan ( – / V \ )
23. wa ( – / V \ )
24. liu ( – / V \ )
25. kui ( – / V \ )
26. tao ( – / V \ )
27. xue ( – / V \ )
28. huan ( – / V \ )
29. tuo ( – / V \ )
30. yang ( – / V \ )
31. tuo ( – / V \ )
32. xi ( – / V \ )
33. pu ( – / V \ )
34. you ( – / V \ )
35. yang ( – / V \ )
36. pu ( – / V \ )
37. you ( – / V \ )
38. wa ( – / V \ )
39. chong ( – / V \ )
40. pu ( – / V \ )
41. kui ( – / V \ )
42. wan ( – / V \ )
43. liu ( – / V \ )
44. you ( – / V \ )
45. jia ( – / V \ )
46. xi ( – / V \ )

47. tao ( – / V \ )
48. pu ( – / V \ )
49. chong ( – / V \ )
50. tuo ( – / V \ )
51. wan ( – / V \ )
52. you ( – / V \ )
53. shen ( – / V \ )
54. tao ( – / V \ )
55. tuo ( – / V \ )
56. wa ( – / V \ )
57. xi ( – / V \ )
58. kui ( – / V \ )
59. tao ( – / V \ )
60. liu ( – / V \ )
61. shi ( – / V \ )
62. chong ( – / V \ )
63. jia ( – / V \ )
64. wan ( – / V \ )

**Part II: Disyllables**

1. xi ( – / V \ ) shi ( – / V \ )
2. huan ( – / V \ ) liu ( – / V \ )
3. pu ( – / V \ ) xue ( – / V \ )
4. jia ( – / V \ ) kui ( – / V \ )
5. chong ( – / V \ ) shi ( – / V \ )
6. tao ( – / V \ ) jia ( – / V \ )
7. xue ( – / V \ ) you ( – / V \ )
8. tuo ( – / V \ ) xi ( – / V \ )
9. wa ( – / V \ ) tuo ( – / V \ )
10. yang ( – / V \ ) wa ( – / V \ )
11. tuo ( – / V \ ) xi ( – / V \ )
12. shen ( – / V \ ) huan ( – / V \ )
13. liu ( – / V \ ) huan ( – / V \ )
14. pu ( – / V \ ) shi ( – / V \ )
15. wa ( – / V \ ) huan ( – / V \ )
16. chong ( – / V \ ) shi ( – / V \ )
17. kui ( – / V \ ) shen ( – / V \ )
18. jia ( – / V \ ) wan ( – / V \ )
19. wan ( – / V \ ) you ( – / V \ )
20. xi ( – / V \ ) xue ( – / V \ )
21. yang ( – / V \ ) xue ( – / V \ )
22. wan ( – / V \ ) wan ( – / V \ )
23. jia ( – / V \ ) kui ( – / V \ )
24. shen ( – / V \ ) tuo ( – / V \ )
25. you ( – / V \ ) liu ( – / V \ )
26. pu ( – / V \ ) chong ( – / V \ )
27. yang ( – / V \ ) wa ( – / V \ )
28. you ( – / V \ ) shen ( – / V \ )
29. tao ( – / V \ ) pu ( – / V \ )

30. kui ( – / V \ ) liu ( – / V \ )
31. chong ( – / V \ ) yang ( – / V \ )
32. tao ( – / V \ ) tao ( – / V \ )

**Appendix C: Tone learning experience questionnaire**

Now that we have finished learning Pinyin, we would like you to share your learning experience with us. Your opinions will be invaluable for us to improve our curriculum and teaching in the future. Thank you very much!

1. According to your learning experience, you feel:

a. The easiest tone to identify is (circle one): 1st, 2nd, 3rd, 4th, toneless

b. The most difficult tone to identify is: (circle one): 1st, 2nd, 3rd, 4th, toneless

c. Do you feel that it is harder to identify tones in disyllable words? (circle one): Yes No.

2. The strategies that you used to help your tone learning (check all that apply)

a. Listen to the audio files of the pinyin exercises in the textbook

b. Repeat after the audio file

c. Compare what I hear with what I say

d. Practice listening and reading of pinyin with my fellow classmate(s) on a regular base, such as meeting one to two times a week.

e. Using available online resources to practice

f. Others: please specify_____

3. Computer-assisted learning experience

a. How do you like the interactive online practice (www.pinyinpractice.com) in helping you get the correct tones? How do you feel while you are doing the exercises?

b. How do you like the way the online exercises are set up? (Either to pick the tones or to supply the missing tones; to get instant feedback; etc.)

c. What improvement do you hope to see of the online exercises?

d. Which one do you prefer: Big group practice in class or on-line practice? Why?

e. Did you use the website, pinyinpractice.com, outside of class? Did you use similar applications to help you get the right tones? If yes, how often?

f. Have you shared the website (www.pinyinpractice.com) with any of your friend? Do you think you will continue to use online applications to help you learn Pinyin? If so, how often do you think you will use them?

g. Have you ever used any other online applications to help you learn a foreign language? If so, what kind of application?

h. The website, Pinyinpractice.com, also has practices for initials and finals. Have you tried those practices? Yes No

If you haven't tried, do you plan to try? Yes No

Do you think we should use the website, Pinyinpractice.com, to practice initials and finals in class? Why or why not?

# Developing Chinese Matching Games: From Inception to Completion
# (创建中文配对游戏)

Chen, Dongdong
(陈东东)
Seton Hall University
(西东大学)
Dongdong.Chen@shu.edu

**Abstract**: There has been a growing interest in incorporating digital games into language instruction. Theoretical and empirical studies on gaming and language development have suggested some pedagogical advantages to using digital games in language teaching and learning. This article addresses the following questions: Can instructors learn to create games for their students? If so, what are the steps to develop digital games that are appealing and educational as well? The focus of this study will be on an experiment that the author recently conducted which resulted in a set of matching games for beginners of Chinese. In describing the process from the inception to the completion of a game, we hope to shed light on the role of gaming in teaching and learning the Chinese language.

**摘要**：语言教育者已逐渐将数码游戏纳入语言教学。这个方面的理论和实证研究表明，数码游戏因其特点有利于语言的教学。本文试图探讨中文教师能否参与以及如何研发教学游戏这两个问题，重点记录作者为初学者创建中文配对游戏的尝试过程。本文介绍作者的研究成果，意在引起中文教师对此领域的关注，起到抛砖引玉的作用。

**Keywords:** Developing digital games, matching games, Chinese language teaching

**关键词**：创建数码游戏、配对游戏、中文教学

## 1. Introduction

In action research conducted on the perception of classroom activities by college learners of Chinese, Chen (2007) found that most of her participants regarded games as more effective and enjoyable than quizzes and homework. This is not surprising because the current generation of students, known as “digital natives,” spends 5,000 to 10,000 hours (out of their total four years at college) playing video games but devote only 5,000

hours to reading books (Prensky, 2001).[1] In a survey of extramural English language-related activities among Swedish youth aged 11 to 12, Sylvén and Sundqvist (2012) found that those children spent more time playing digital games than watching TV or listening to music. No doubt, games consume a significant portion of time in students' daily life. Interestingly, researchers found that "gamers" play various types of games throughout different stages of their lives (e.g., Götzenbrucker & Eludamos, 2009; Quandt, Grueninger, & Wimmer, 2009; Chik, 2014). Takahashi (2015) projected that the video gaming industry will increase by 30% from 2010 to 2019, reaching $19.6 billion in revenue, suggesting that the popularity of these games will continue to increase.[2] While some games are "designed purely for fun and entertainment" (Whitton, 2010, p. 199), many have been utilized in education by professionals of various disciplines, including teaching and learning world languages (e.g., Caillois, 1961; Gee, 2007; Reinhardt, 2012; McGonigal, 2013; Sykes & Reinhardt, 2013; Sykes, 2018).

What is a game? Why is its appeal so powerful that serious educators have turned their attention to it as a potential instructional tool? Pearce (2002) states that a game is a structured framework that consists of: a goal; obstacles; resources; rewards; penalties, and; information. Khine (2011) defines the character of game playing as consisting of "rules, goals, engagement, challenge, feedback, fun, interactive, outcome, and immediate reward" (p. 121). In his book *Language Play*, *Language Learning,* Guy Cook, an applied linguist, distinguished "games" from "play," though he noted that these two words are interchangeable in many cases. According to Cook (2000), "game" has two different senses. In general, "'game' is used to describe a range of activities…" Specifically, a game "refers to intricate, rule-governed, and culturally variable competitive activities" (p. 127). While "play" is a "free activity" and "rule-governed," it is "not serious" (p. 112). As structured activities, both "game" and "play" can engage learners to study. A digital game is designed to be played on an electronic device, such as a computer, smart phone, iPad, or videogame console. As Khatibi and Cowie (2013) remarked that, "games, whether they are digital or not, embody some specific features that facilitate learning… Digital games also provide the learners with a platform that supports interaction..." (p. 35).

Given that learners in the 21st century are exposed to a wide variety of emerging and evolving technologies, the teaching approaches, methods, and strategies that performed well for students of former generations might not be as effective for this generation. Hence, some researchers suggest a change of pedagogy. For example, Prensky (2007) proposed incorporating technology into the classroom to accommodate learners with different backgrounds and needs. Khatibi and Cowie (2013) pointed out that to help today's students learn, educators must connect to them by acknowledging that digital games are a significant part of their lives. Godwin-Jones (2014) remarked that integrating gaming into language learning is a "winning situation for both students and educator" (p. 9). Based on the research in the field, Sykes (2018) concluded that the

[1] They spend 10,000 hours speaking on cell phones, 20,000 hours watching TV, and 250,000 hours sending messages.

[2] In Duggan's (2015) survey of 2001 Americans, 49% of them reported about playing games and 10% considered themselves "gamers."

"incorporation of digital games into world language teaching and learning offers interesting and varied possibilities" (p. 220).

Over the past ten years, numerous theoretical and empirical studies have been conducted on digital gaming and language development (e.g., Pomerantz & Bell, 2007; Thorne, 2008; Thornet et al., 2009; Talak-Kiryk, 2010; Benson & Chik, 2011; Reinders, 2012; Peterson, 2013; Khatibi & Cowie, 2013; Godwin-Jones, 2014; Lan, Fang, Legault, & Li, 2015; Ketterlinus, 2017; Sykes, 2018). These studies have found many pedagogical advantages of digital games in language teaching and learning. Pomerantz and Bell (2007), for instance, observed that gaming introduces "fun" and "creativity" to the language classroom. Gameplay motivates learner autonomy which, in turn, leads to the retention of what is being learned. Prensky (2007) argued that the reasons why a digital game facilitates learning are its features of engagement, interactivity, and the combination of the two aspects. Sylvén and Sundqvist (2012), in the previously-mentioned study of the young Swedish subjects, discovered that frequent gamers (those playing games 5 or more hours a week) achieved the highest scores on English proficiency tests, followed by moderate gamers and non-gamers. The researchers concluded there was a positive correlation between L2 gaming and L2 learning. Godwin-Jones (2014) noted three benefits. First, gamers, while participating in a "massively multiplayer online game," are inspired to use the target language actively in a socially appropriate context. Secondly, as the gaming system provides continuous feedback, players are encouraged to repeat, revise, and reproduce constantly. Finally, players enjoy playing because of a sense of accomplishment.

Despite the extensive literature of gaming and world language education, few studies develop Chinese games for learners or examine the learning of Chinese in a gaming context. Yao and McGinnis (2002) and Chen and Fellows (2019) are two notable exceptions. They create games or game-like communicative activities for the Chinese classroom. Lan, Fang, Legault, & Li (2015) investigated how the Virtual Environment (VE) impacted the L2 acquisition of Chinese vocabulary and showed that the VE was able to accelerate the learning of vocabulary. As L2 gaming has been shown to facilitate L2 teaching and even enhance L2 learning in some aspects, it is appropriate that Chinese language teachers examine some questions in this regard. For example, are there any commercial digital games that are useful in the Chinese classroom? If so, what needs to be done to integrate them into the curriculum? Since instructors know their students better than anyone else, can they learn to create games for their students? If the answer is "yes," what kind of games should they create? What are the steps to develop digital games that are appealing and educational as well? This article attempts to address some of these questions. The focus will be on an experiment that the author recently conducted which resulted in a set of matching games for beginners of Chinese. In describing the process from the inception to the completion of a game, we hope to shed light on the role of gaming in teaching and learning Chinese.

## 2. Why Gaming for Learning Chinese?

Learning a language requires a strong commitment. However, it rarely yields an immediate satisfactory outcome or instant gratification. The process is time-consuming and challenging. The learners' motivation and responsibility are among the many factors that contribute to a successful experience. A good learner holds an enduring interest in the language, keeps practicing, endeavors to use the language whenever and wherever he or she can, and strives for communicative competency. Such an active attitude and responsible learning behaviors are valued by those who advocate communicative language teaching (CLT) (e.g., Hymes, 1972; Krashen, 1982; Howatt, 1984; K. Johnson & J. Johnson, 1998; Spada, 2014). Under the philosophy of CLT, the instructor is just a "facilitator" or "planner" who engages learners to study by offering them a supportive learning environment (Richards & Rodgers, 2001). Instead of feeding students knowledge about grammar, the instructor formulates meaningful tasks to maximize learners' use of the language. Students thus become active participants in the process by interacting with others. In this way, students learn to communicate in the target language. Student-centered learning and the meaningful use of instructional materials are the essence of CLT.

Games for language learning, as argued by Sykes and Reinhardt (2012), possess some inherently important features. First, games have a learner-directed goal orientation, which drives learners towards their own objectives. Secondly, games provide plenty of opportunities for learners to interact with games and with other gamers while playing. Third, individualized and timely feedback encourages learners to constantly improve. Fourth, games create a context that provides a meaningful experience for learners. Finally, learners are motivated to play and enjoy playing because it provides an engaging experience. Obviously, these characteristics resemble the attributes promoted by CLT and reflect best practices in second language teaching and learning. Lee (2016) described the benefits of some game-like learning principles related to gaming.[3] When one is involved in playing games, he or she is a participant. Learning happens by doing. Games provide ongoing challenges with constant and helpful feedback. If a student fails, the failure is reframed in a new iteration. In this process, learning feels like play, with no stress or struggle, hence reducing affective filters (Krashen, 1982).

Researchers in the field distinguish game-enhanced learning from the game-based learning (e.g., Godwin-Jones, 2014; Sykes, 2018).[4] The former refers to the use of commercial, off-the-shelf games for learning while the latter includes the use of digital games designed explicitly for the teaching and learning of world languages. Even though the international Edutainment (i.e., *education* and *entertainment*) Conference was initiated in China in 2006, and has been held there a few times since, almost nothing has been published on game-supported learning with respect to teaching Chinese as a second

[3] Gee (2007) identified 36 learning principles present in many of the games.

[4] Sigurðardóttir (2010) classified 11 educational games: games for dividing larger groups into smaller groups, introduction-games, group games, physical games, scavenger hunt games, educational games, theoretical expression games, drawing and coloring games, educational card games, word games, story games, and question games.

language.[5] This is puzzling. But if we take into consideration Chinese views of the word "game," we might understand this situation. When the sense of "game" in the context of learning is translated into Chinese, it is *yóuxì* 游戏, meaning "play." Chinese renditions for "play," "pastime," and "play games" are all *yóuxì* 游戏. In the 6th edition of *Modern Chinese Dictionary* (Institute of Linguistics, 2012), *yóuxì* 游戏, when used as a noun, means *entertainment*, and refers to *play* when it is a verb. The English phrase "play games" is translated as "*wán yóuxì* 玩游戏." To many Chinese, parents in particular, *xuéxí* 学习 [learn] or yánjiū 研究 [study] is serious and painstaking work, while *yóuxì* 游戏 is not. Furthermore, the word *yóuxì* 游戏 has some derogative connotations. Playing games in kindergarten is fine, but pupils would be discouraged or criticized for doing so in elementary school. In fact, buried among much homework and assignments in preparation for numerous tests, Chinese students are deprived of opportunities for play. Students are expected to work diligently on school subjects, while instructors are supposed to teach wholeheartedly. As such, serious teachers would require students to spend all their school time and beyond "working" on subjects rather than "playing" educational games. These cultural assumptions were captured in Chik's (2012) study, which examined the perspectives of students and teachers in Hong Kong about digital gameplay for autonomous foreign language learning. Chik (2012) found that teachers considered gameplay as an independent activity, but not connected with language learning and use in gamers' personal and social worlds. Most of the teachers in the study thought that English learning through gaming was not possible for most gamers. Some teachers disapproved of online gaming, considering it a waste of time.

Further research should be conducted to explore attitudes towards the use of games in education.[6] This article examines how digital games can help meet the pedagogical needs of learning Chinese. As is commonly known, Chinese is one of the most difficult languages for English speaking learners due to tones or the writing system or both (e.g., DeFrancis, 1984). The difficulty with tones is self-explanatory because of tonal features. From learning perspectives, learners must first perceive the subtle nuance of each of the four different pitch patterns. The level tone (or the 1st tone), the rising tone (or the 2nd tone), the dipping tone (or the 3rd tone), and the falling tone (or the 4th tone) take time to get used to for most beginners, particularly so for those who are not "musical." When reading *pinyin*, a Chinese Romanization system, learners must identify the four tonal representations, and determine which diacritic stands for which tone. Marked by the four diacritics, ˉ, ˊ, ˇ, ˋ, placed on the top of the primary final (i.e., vowel), these tone graphs which look simple may appear counter-intuitive to learners, as argued by Bar-Lev (1991). According to Bar-Lev (1991, p. 8), "Most Americans can no more easily remember diacritics than numbers of the tones themselves." This suggests that the conventional method to represent Chinese tones is as confusing and perplexing. Not only does it fail to alert the learner that each of the tone accents is an essential part of a sound unit, the tone graph is not meaningful enough to remind the learner that each syllable carries a unique tone. In other words, what learners see does not match what they

[5] See Zhang et al.'s (2010) *Proceedings of the 5th International Conference on E-learning and Games*.

[6] Whitton (2012) reported some problems of the acceptability of games in formal educational contexts. Also see Duggan (2015) for public debates about "gaming and gamers."

hear. Furthermore, unlike English, which seems to have "tone" or rather "intonation" at the sentence level, for example, a rising tone for posing a question or a falling tone for making a statement, Chinese tones are lexically localized. Thus, when it comes to the production of tones, someone speaking English only needs to raise or lower one's pitch at the end of a sentence. However, to speak Chinese one must twist his or her tongue for almost each word.

In terms of the difficulty with the Chinese writing system, Chinese characters are composed of radicals rather than an alphabet. A radical is a basic element that has its unique image, sound, and meaning. In most cases, a radical is a character itself. Sometimes a radical may have a variant. In that case, it cannot appear alone as a character, as shown by "亻"—a form of the radical rén 人 [people]. Most Chinese characters are made of two or more radicals. Thus, the journey of learning to read and write Chinese starts with each individual radical. In addition, being logographic in nature, Chinese writing does not carry any association between how a syllable is pronounced (i.e., the *pinyin* form) and how it is written (i.e., character). Thus, to learn each radical or character, one must memorize three elements simultaneously: image, pronunciation, and meaning.

Most beginners find Chinese tones and characters difficult. Requiring students to learn both at the same time is not effective. DeFrancis (1966) once proposed diagraphia to deal with such an issue. That is, at the beginning stage, the instructor only focuses on pronunciation and speaking. Only when the learners feel comfortable about sounds and tones, does the instructor start to introduce the Chinese writing. This proposal addresses one objective at a time, allocating a period of time to learning the *pinyin* system before students start to learn characters. To strengthen the learner's aural and oral skills, DeFrancis compiled his textbooks in two different versions: one in the form of *pinyin* and one in the form of characters. This solution has proved helpful to learners, as his 12-volume "DeFrancis Series" became the most widely used teaching resources in the 1970s and 1980s (e.g., Chen, 2013). Packard (1990) compared the learning outcome of a "lag group" (i.e., studying *pinyin* exclusively for three weeks before the introduction of character) with that of the "no-lag group" (i.e., studying *pinyin* and character at the same time in the first week). The findings showed that the students in the lag group were significantly better than those in the no-lag group in phonetic discrimination, and they were significantly more fluent in oral production.

Regardless of the length of time devoted to the learning of *pinyin* before the introduction of characters, the two elements are challenging for beginners. It is also obvious that the rote learning of these forms is absolutely required. For learners of Chinese to reach an ultimate proficiency in the language, they must experience numerous repetitions of different tones, various characters, new words, special collections, linguistic patterns, etc. The question then is whether the current "digital natives" or "game generations" (Prensky, 2001; Prensky, 2007) can endure such a long and tedious process. It would be difficult to retain their enthusiasm because of boring repetition unless they are extremely motivated or disciplined learners. Digital games address these learning issues. This is because when playing games, learners are fully engaged in achieving the goals without consciously realizing all the manual and mental repetitions,

thereby not only reducing boredom and tediousness, but increasing effectiveness and efficiency. Along the lines of remark by Godwin-Jones (2014) "gameplay should not be introduced in a language classroom without an awareness of the practical, pedagogical, and personal issues involved" (p. 14), we argue that the learning issues of Chinese tones and characters are rationales for the incorporation of games into the teaching process. If there are no such digital games to be found that address the issues, then let us create one.

## 3. What Games Are Appropriate?

As illustrated above, whether it is to learn tones or characters, beginners must be able to identify an oral or visual form and produce them. Assuming that learning Chinese implies a transformation from not knowing how to pronounce a word to being able to produce sentences, from struggling with tones and characters to being able to use them comfortably, we believe that the fundamental activity is memorization. Bilbrough (2011) claimed, "There is no learning without remembering. And language learning—perhaps more than most forms of learning—places huge demands on memory" (p. 1). The memory skill, argued by Bilbrough to be the fifth skill of language learning, is vital for the development of the skills of listening, speaking, reading, and writing. Remembering language involves encoding, storage, and retrieval, supported by certain parts of the brain. In their research that examined the neuroscientific mechanisms involved in language learning and gaming, Khatibi and Cowie (2013) found that both share certain aspects of the same processes. They deduced that "the memory of previously encountered situations would facilitate both processes since what we remember serves as a foundation for learning new information simultaneously" (p. 26). This implies that a game for building memory would benefit the learning of tones and characters. Therefore, we decided to create a means for learners to perform matching exercises to reinforce the associations among different elements of a given character, including its pronunciation with a correct tone, its meaning, and its image. Based on such an understanding, eight matching pairs were conceived, as shown in Table 1.

**Table 1 8 Pairs of Matchings**

| **Category** | **Example** |
|---|---|
| Matching *Pinyin* to English | rén ⟷ person |
| Matching Character to English | 人 ⟷ person |
| Matching *Pinyin* to Character | rén ⟷ 人 |
| Matching *Pinyin* to Sound | rén ⟷ 🔊 |
| Matching Character to Sound | 人 ⟷ 🔊 |
| Matching *Pinyin* to Tone Graph | rén ⟷ ↑(35) |
| Matching Character to Tone Graph | 人 ⟷ ↑(35) |
| Matching Radical to Radical | 人 ⟷ 人 |

Except for the last one, all the above seven categories have one thing in common: enabling learners to build an association between the two components. The repeated exposure to these elements throughout the matching game will help learners to retain

these connections, which eventually become a part of their own mental repertories. The last pair is a combination of two radicals forming a single, new character of which there are two models. The models can be either a semantic-semantic compound or a semantic-phonetic compound, representing two productive ways to form Chinese characters (e.g., Sun, 2006). These eight matchings encourage and push learners to develop their memory skills, which will help them recognize, store, retrieve, and produce tones and characters.

## 4. How to Develop Games?

With the general principles of the matching game outlined in the above section, the next question is: how to implement the design? Another immediate and more crucial question is: which tool to use for creating the matching game? The criteria of such a tool should be simplicity and ease of use. That is, the tool must be free, easy to learn, but have a wide range of practical features. Along these guidelines, UNITY was selected as the tool.[7] As an open-source tool, UNITY is a leader in the global game industry. While it is claimed to be easy, UNITY involves coding in C#, which is a daunting task for language teachers who are not trained in programing and do not have time to do so. After taking three tutorials about the basics of UNITY, the author concluded that it would be unrealistic to learn how to code in C# first and then to implement the matching game. In order to produce a game-like effect for the anticipated game design while maintaining the pedagogical goal, the author decided to adopt a practical approach—to purchase a commercial product instead of creating one and then to modify it by inserting instructional material. Such a short-cut, providing it is legally permissible and technically feasible, would greatly facilitate the development of the game.[8]

After many searches in the UNITY Asset Store, the author located the Matching Game Template created by Puppeteer (Abdulqadir, 2017) and purchased it. This template is an action-packed game that offers attractive features: 1) it supports different operating systems, like PC/Mac, iOS, and Android; 2) it is easy to customize, and; 3) it includes all the key assets, such as graphics, sounds, and the source code. The template also includes user-friendly documentation. Above all, as a certified UNITY developer specializing in creating game templates, the builder "aims to produce high-quality packages with complete functionality to give other developers a head-start" (Abdulqadir, 2017). Online support is also available.

In developing the Chinese matching game using this template, it was important that the basic mechanism of the template be understood. By playing the game and studying the documentation, the author came to understand the logic of the game. However, as a novice to UNITY and game creation, the author could not determine how to insert Chinese characters into the gaming system. Without resolving this problem the

[7] c.f. http:// www.unity.com. The author is most grateful for Veronica Armour's recommendation of this tool.

[8] See Reinders and Wattana (2012) for modifying an existing game, *Ragnarok Online*, the most popular massively multiplayer online role-playing game in Thailand, for educational purposes. Also, see Vandercruysse et al. (2013). Again, I thank Veronica Armour for her help.

gaming development could not proceed. The author thus contacted the owner of the template for help. Amazingly, the owner immediately provided the help that was needed. With that initial assistance and considerable consequent support, and with many hours of hard labor, the author was able to produce the first version of a set of three matchings, followed by a second version of four matchings. Eventually, a third version of all the eight matchings was completed. The first and the third versions of the games are available for play on the site http://tltc.shu.edu/chinese/Game/. They work with either FireFox or Chrome browsers. The second version of the game was created for Android-based mobile phones.[9] The interface of the game is shown in Fig. 1. Table 2 and Table 3 list the screenshots of the eight possible matchings.

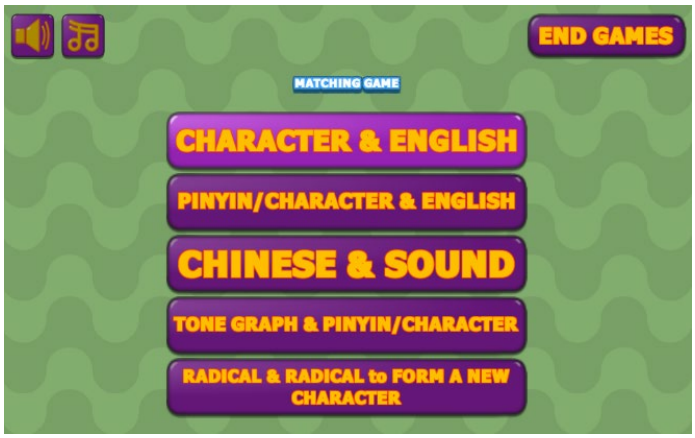



**Figure 1 Interface of the Matching Game**

As shown in Fig. 1, a player can click any of the five categories to start the game. The function of each object is self-explanatory.

**Table 2: The Screenshots of the Matching Pairs:**
**Chinese to English; Tone to *Pinyin*/Character; Character to Character**

| Categories | Examples |
|---|---|
| **Matching *Pinyin* to English** |   |
| **Matching Character to English** |   |
| | |

[9] The author is willing to provide the APK for users of Android phones upon request.

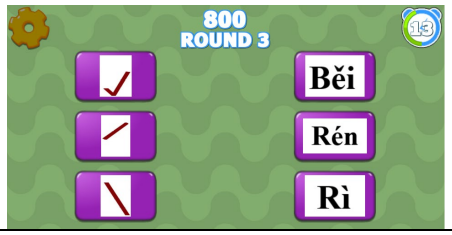
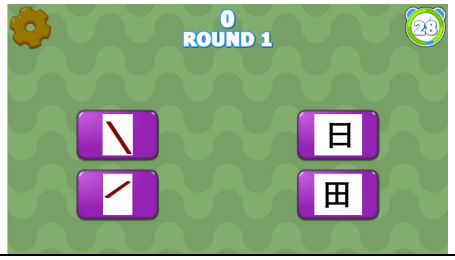
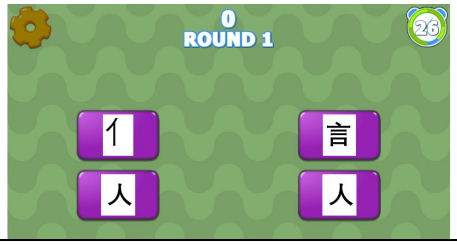
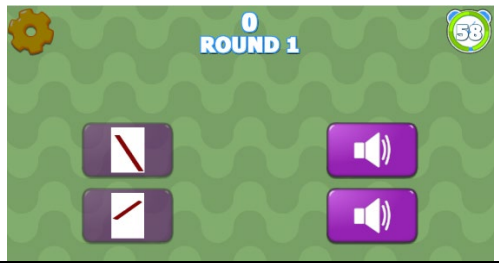
| | |
|---|---|
| **Matching *Pinyin* to Character** |  |
| **Matching Tone Graph to *Pinyin*** |  |
| **Matching Tone Graph to Character** |  |
| **Matching Radical to Radical** |  |

**Table 3: The Screenshots of Matching Pairs: Tone Graph to Sound; Character to Sound**

| Categories | Examples |
|---|---|
| **Matching Tone Graph to Sound** |  |
| **Matching Character to Sound** |  |

A few words are in order to describe how the game works. Clicking "Character & English," for example, the user will be presented with two pairs of characters and English words. If the matching of the first pair is correct, that pair will disappear, and 100 points will be awarded, shown on the top middle. If not correct, the two items remain on the screen. When the two pairs are finally matched, the player advances to the second level, which has three pairs. For each following level, one more pair is added, while the time is reduced, as shown by the timer icon on the top right. As the level of difficulty increases with more pairs popping up to be matched, creating complexity and confusion, so do the award credits, with 100 more points added for each matched pair in the higher level. When the player matches all the pairs within the given time, he or she wins the game! This makes the task challenging and fun.

This matching game, simple as it may be, fully reflects the major characteristics that a game should have, as defined by Pearce (2002). Winning the game is the goal. Each extra pair added with less time creates obstacles, preventing the player from reaching the goal. When a pair is matched, the reward is a score of 100 points. This further stimulates and motivates the player to work harder so as to achieve ultimate victory. When time runs out and there are still pairs to match, the game is over, resulting in penalties. Most players will likely repeat the game in hopes of performing better. Through repeated trials with feedback, fun and interactive engagement leads to learning. Each matching, correct or wrong, is an informative and meaningful experience, helping the player to build and strengthen the association between two related items.

## 5. Discussion and Conclusion

Creating a digital game is a major undertaking for language teachers, but it is achievable. Throughout this experiment, the author learned valuable lessons. Building a game involves three steps. First, determine the pedagogical goals, because they justify the needs of introducing games to the classroom. What are the linguistic phenomena that usually cause problems for learners of the language? Why are those elements difficult? Are they related to each other? What can a game do to support learning? Identifying these learning issues is the first step. By clarifying these concerns, consider the most appropriate format of the game for different learning targets. When designing games, focus on the user's experience. Endeavor to make the game stick to the learning goal, not just for the sake of fun, because gaming, like technology, should be a means to help achieve what could not be achieved as effectively by other methods (e.g., Oxford & Oxford, 2009). Having designed the game, proceed with its implementation. A good developing team should consist of: 1) a savvy developer who specializes in game creation; 2) a well-informed instructional designer who knows educational technologies, and; 3) an experienced language teacher who understands pedagogical requirements. Joint efforts from these experts are ideal but not always feasible. A practical recommendation from the author is to purchase a game and adapt it for educational purposes. During the process of gaming development, work closely with the builder of the game template and discuss issues with the instructional designer regularly for feedback and suggestions. The current experiment is in accordance with the spirit of approaches that Whitton (2012) proposed to

solve the problem of cost related to the purchase of commercial games and to the time needed for educators to attain the skills to develop games. One approach involves educators trying out free or inexpensive game development toolkits and techniques to create games. For educators who are interested in developing matching games, the author recommends the template as described in this article.

Through the experience gained undertaking this experiment, the author plans to develop more games. There are two directions in this regard. One is to engage the students of Chinese to create more matching games based on the current template. A pilot study of this will be reported in another article.[10] The other is to create new games of different types. To that end, Chinese characters remain one of the primary focuses. For example, one possible game might require users to search for a correct character from a pile of characters based on a given pronunciation, visual image, meaning, or a combination of the three. This kind of identification game will help reinforce learners' familiarity with characters. Another potential game would involve Chinese grammatical points that are confusing to learners. For instance, prepositional phrases such as "shàngkè yǐqián" 上课以前 [before class], "xiàkè yǐhòu"下课以后 [after class] always pose learning problems. It is common to see learners' utterances like "yǐqián shàngkè" 以前上课, and "yǐhòu xiàkè"以后下课. This is because these prepositions, yǐqián 以前 [before] and yǐhòu 以后 [after], are head finals in Chinese, which is the opposite word order of their English counterparts. Related to this phenomenon is the Chinese structure where the locative phrase usually occurs before the verb, such as "zài túshū guǎn gōngzuò" 在图书馆工作 [work at the library], whereas for certain verbs the locative may appear post-verbally, e.g., "tǎng zài chuángshàng" 躺在床上 [lie in bed]. Chinese relative clauses require the modifier preceding the nominal, different from English, which has a reverse word order regarding the two elements. All these challenges would make good material for a game on word order. Chinese idioms would be another interesting candidate, as they have a rigid word order and wording, which learners often find difficult.

Having succeeded in creating Chinese matching games, the author observed that these games have three major benefits for the teaching and learning of Chinese. First, students can play the games to learn and review material. Available online, players can access the games anywhere and at any time. Second, the instructor can use the games as teaching resources in the classroom to engage student learning. The third benefit is that the games can be recycled with different content. For example, the author initiated a pilot study in the spring of 2018 in which students enrolled in the second part of Introductory Chinese were able to use a template to create additional matching games.

[10] The author initiated a pilot study in the spring of 2018 in which students enrolled in the second part of Introductory Chinese created more matching games. Working in teams of two, each group was required to produce a set of matching games for the vocabulary of one chapter of *Integrated Chinese, Level 1, Part 1*. The students were provided with the game's template and step-by-step instructions. The students ultimately produced six sets of matching games covering six chapters of the textbook.

To conclude, the author anticipates conducting empirical studies in the future to compare the learning outcomes of Chinese tones and characters via the gamed-supported learning versus traditional learning. Findings of this type of research will provide insight to those who have been exploring the value of gaming in learning world languages and the teaching of Chinese in particular. Positive results of the effectiveness of game-assisted teaching may inspire Chinese language teachers either to seriously consider introducing games in their classroom or to undertake gaming development to empower their teaching.

**Acknowledgement**: I acknowledge the guidance and support of Veronica Armour, Instructional Designer, Thomas A. McGee, Senior Digital Media Specialist, and Obed E. Raimundo, Web developer of the Teaching, Learning and Technology Center, Seton Hall University. Special thanks go to Mary M. Balkun and Marta M. Deyrup for their encouragement. I am grateful to the University DH Fellowship that I received in 2017, which made this project possible. Part of this work was presented at the Association of Taiwanese Teachers in Mainstream Schools in NY/NJ, New Jersey, on January 21, 2018, and the 4th Symposium on Teaching Chinese as a Second Language held at George Washington University, Washington D.C. on May 7, 2018. I thank the audiences for constructive comments and questions.

## References

Abdulqadir, M. (2017). Matching game template. Retrieved from https://connect.unity.com/u/5818ce8f090915002eeb13bc

Bar-Lev, Z. (1991). Two innovations for teaching tones. *Journal of Chinese Language Teachers Association*, *26*(3), 1-24.

Benson, P., & Chik, A. (2011). Towards a more naturalistic CALL: Video gaming and language learning. *International Journal of Computer-Assisted Language Learning and Teaching*, *1*(3), 1-13.

Bilbrough, N. (2011) *Memory activities for language learning*. Cambridge: Cambridge University Press.

Caillois, R. (1961). *Man, play, and games*. Glencoe, IL: The Free Press.

Chen, D. (2007). Games: Why are they popular among CFL students? In Y. Zhu & T. Yao (Eds.), *Proceedings of the 5th international conference on Chinese language pedagogy* (pp. 262 -269). Beijing: World Publishing Corporation Press.

Chen, D. (2013). 汉语在美国的传播历史 -- 纪念美国中文教学界三位巨匠 (History of teaching Chinese in America----In commemoration of three giants). 《汉语国际传播研究》(*The Journal of Chinese Language Globalization Studies*), 2, 16-27.

Chen, D. (2013). History of teaching Chinese in America----In commemoration of three giants. *The Journal of Chinese Language Globalization Studies*, 2, 16-27. [陈东东. 2013. 汉语在美国的传播历史 -- 纪念美国中文教学界三位巨匠. *汉语国际传播研究*, 2, 16-27.]

Chen, D., & Fellows, C. (2019). *50 activities for the international Chinese classroom*. Beijing: Foreign Language Teaching and Research Press.

Chik, A. (2012). Digital gameplay for autonomous foreign language learning: Gamers' and language teachers' perspectives. In H. Reinders (Ed.), *Digital games in language learning and teaching* (pp. 95-114). London: Palgrave Macmillan.

Chik, A. (2014). Digital gaming and language learning: Autonomy and community,

*Language Learning & Technology*, *18*(2), 85-100. Retrieved from https://www.lltjournal.org/item/2857.

Cook, G. (2000). *Language play, language learning*. Oxford: Oxford University Press.

DeFrancis, J. (1966) Why Johnny can't read Chinese? *Journal of the Chinese Language Teachers Association*, *1*(1), 1-20.

DeFrancis, J. (1984). *The Chinese language: Fact and fantasy.* Honolulu, HI: University of Hawaii Press.

Duggan, M. (2015). Games and gamers. Pew Research Center—Internet and Technology. Retrieved from http://www.pewinternet.org/2015/12/15/gaming-and-gamers/

Gee, J. P. (2007). *What video games have to teach us about learning and literacy*. New York: Palgrave Macmillan.

Godwin-Jones, R. (2014). Emerging technologies games in language learning: Opportunities and challenges, *Language Learning & Technology*, *18*(2), 9-19. Retrieved from https://www.lltjournal.org/item/2849

Götzenbrucker, G. & Eludamos, M. (2009). Ten years later. Towards the careers of long-term gamers in Austria. Eludamos. *Journal for Computer Game Culture*, *3*(2), 309-323. Retrieved from http://www.eludamos.org/index.php/eludamos/article/viewArticle/74

Howatt, A. P. R. (1984). *A history of English language teaching*. Oxford: Oxford University Press.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics* (pp. 269-293). Harmondsworth: Penguin.

Institute of Linguistics, Chinese Academy of Social Sciences. (2012). *Modern Chinese dictionary*. Beijing: The Commercial Press. [中国社会科学院语言研究所词典编辑室]. (2012). *现代汉语词典*. 北京: 商务印书馆.]

Ketterlinus, L. (2017). Using games in teaching foreign languages. MA Thesis, United States Military Academy, West Point, NY. Retrieved from https://www.usma.edu/cfe/Literature/Ketterlinus_17.pdf

Khatibi, E. & Cowie, E. (2013). Language learning through interactive games. Retrieved from http://muep.mau.se/handle/2043/16116

Khine, M. S. (Eds.) (2011). *Playful teaching, learning games: New tools for digital classrooms*. Rotterdam: Sense Publishers.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.

Johnson, K., & Johnson, J. (1998). Communicative methodology. In K. Johnson & H. Johnson (Eds.), *Encyclopedic dictionary of applied linguistics* (pp. 68-73). Oxford: Blackwell.

Lan, Y., Fang, S., Legault, J., & Li, P. (2015). Second language acquisition of Mandarin Chinese vocabulary: context of learning effects, *Educational Technology Research and Development*, *63*(5), 671-690.

Lee, J. L. (2016, March). Games for social and educational impact: Lessons learned and what's next? Gamification Symposium, Rutgers University, New Brunswick.

McGonigal, J. (2013). *Reality is broken: Why games make us better and how they can change the world*. New York: The Penguin Group.

Oxford, R., & Oxford, J. (Eds.). (2009). *Second language teaching and learning in the net generation* (pp. 1-8). National Foreign Language Resource Center, University of Hawaii.

Packard, J. L. (1990). Effects of time lag in the introduction of characters into the Chinese language curriculum. *Modern Language Journal, 74*(2), 167-175.

Pearce, C. (2002). Towards a game theory of game. In L. King (Ed.), *Game on: The history and culture of video games* (pp. 112-119). London: Lawrence King.

Peterson, M. (2013). *Computer games and language learning.* New York: Palgrave Macmillan.

Pomerantz, A. & Bell, N. (2007). Learning to play, playing to learn: FL learners as multicompetent language users. *Applied Linguistics*, *28*(4), 556-578.

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon*, *9*(5), 1-4.

Prensky, M. (2007). *Digital game-based learning*. St. Paul, MN: Paragon House.

Quandt, T., Grueninger, H., & Wimmer, J. (2009). The gray haired gaming generation: Findings from an explorative interview study on older computer gamers. *Games and Culture*, *4*(1), 27-46.

Reinhardt, J. (2012). *Digital games in language learning and teaching*. London, UK: Palgrave Macmillan.

Reinders, H., & Wattana, S. (2012). Talk to me! Games and students' willingness to communicate. In H. Reinders (Ed.), *Digital games in language learning and teaching* (pp. 156-188). Basingstoke, UK: Palgrave Macmillan.

Richards, J., & Rodgers, T. (2001). *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.

Sigurðardóttir, S. D. (2010). The use of games in the language classroom, University of Iceland. Retrieved from http://hdl.handle.net/1946/6467.

Spada, N. (2014). Research on second/foreign language learning: Some implications for teaching, *The Journal of Chinese Teaching and Research in the US*, *9,* 1-16.

Sun, C. (2006). *Chinese: A linguistic introduction*. Cambridge: Cambridge University Press.

Sykes, J., & Reinhardt, J. (2012). Language at play: Digital games in second and foreign language teaching and learning. In J. Liskin-Gasparro & M. Lacorte (Series Eds.), *Theory and practice in second language classroom instruction* (pp. 1-157). New York: Pearson-Prentice Hall.

Sykes, J. M., & Reinhardt, J. (2013). *Language at play: Digital games in second and foreign language teaching and learning*. Boston, MA: Pearson.

Sykes J. M. (2018). Digital games and language teaching and learning. *Foreign Language Annals*, *51*, 219-224. Retrieved from https://doi.org/10.1111/flan.12325

Sylvén, L. K., & Sundqvist, P. (2012). Gaming as extramural English L2 learning and L2 proficiency among young learners. *ReCALL*, *24*(3), 302-321.

Takahashi, D. (2015). U.S. games industry forecast to grow 30 percent to $19.6B by 2019.*Venture beat*. Retrieved from https://venturebeat.com/2015/06/02/u-s-games-industry-forecast-to-grow-30-to-19-6b-by-2019/

Talak-Kiryk, A. (2010). Using games in a foreign language classroom. MA TESOL Collection. Retrieved from https://digitalcollections.sit.edu/ipp_collection/484

Thorne, S. L. (2008). Transcultural communication in open Internet environments and massively multiplayer online games. In S. Magnan (Ed.), *Mediating discourse online* (pp. 305-327). Amsterdam: John Benjamins.

Thorne, S. L., Black, R. W., & Sykes, J. (2009). Second language use, socialization, and learning in Internet interest communities and online games. *Modern Language Journal*, *93*, 802-821.

Vandercruysse, S., Vandewaetere, M., Cornillie, F. & Clarebout, G. (2013). Competition and students' perceptions in a game-based language learning environment. *Education Tech Research*, *61*, 927-950.

Whitton, N. (2010). *Learning with digital games: A practical guide to engaging students in higher education*. New York: Routledge.

Whitton, N. (2012). The place of game-based learning in an age of austerity. *Electronic Journal of e- Learning, 10* (2), 249-256

Yao, T. & McGinnis, S. (2002). *Let's play games in Chinese*. Boston: Cheng & Tsui Company.

Zhang, X., Zhong, S., Pan, Z., Wong, K., & Yun, R. (Eds.) (2010). *Proceedings of the Entertainment for education. Digital techniques and systems.* 5th International Conference on E-learning and Games, Edutainment 2010, Changchun, China, August 16-18, 2010.

http://www.tclt.us/journal/2019v10n1/bailiyeh.pdf pp. 73-90

# Integrating Technology in the Teaching of Advanced Chinese (高年级中文课教学的技术应用)

Bai, Jianhua
(白建华)
Kenyon College
(肯扬大学)
bai@kenyon.edu

Li, Cong
(李聪)
Duke Kunshan University
(昆山杜克大学)
li.2908@osu.edu

Yeh, Wen-Chin
(葉雯瑾)
Kenyon College
(肯扬大学)
wenjin.yeh@gmail.com

**Abstract:** This article deals with the rationale, design, and implementation of an advanced Chinese course delivered in a distance-learning environment. The course helps students of advanced Chinese develop their language competency and independent research skills through an instructional model that consists of individually directed research, small group discussions, one-on-one practice sessions, and individual presentations followed by peer critiques and Q&A sessions. Various technology tools, such as Zoom, VoiceThread, and an online corpus, are employed to facilitate the entire process of teaching and learning in this type of distance learning course. Based on student feedback and our evaluation of the course at the end of the semester, we feel very positive about the addition of this new course to our permanent Chinese language curriculum.

**摘要：**本文旨在探索远程环境下高级中文独立研究课程之设计以及科技辅助的成效。课程主体依循一套学生独立研究-面谈-演练-报告的基本模式，通过 Zoom，VoiceThread 和语料库等科技辅助手段，优化远程教学及互动经验。融合了线上同步个人面谈、线上非同步个人报告与同侪反馈、团体论坛，以及线下辅导课。强调个人独立学习与团体合作，以提升学生语言表达能力，并帮助学生提高批判性思维和自主学习的能力。本文亦讨论这种新形态汉语课程的优势，根据学生的反馈和我们在教学中的反思及学期末教学成果的评估，我们计划将这门新形态独立研究课程编入我们的高级中文系列教程。

**Keywords**: Distance learning, advanced Chinese, independent research, technology tools

**关键词:** 远程教学、高级中文、独立研究、科技辅助

## 1. Introduction

The field of teaching and learning Chinese as a foreign language (CFL) has come a long way (Bai, 2007; Ling, 2018) and great changes have occurred in the past two decades. We have witnessed a growing interest in Chinese language learning at all levels, from elementary and secondary schools to institutions of higher education (Goldberg, Looney, & Lusin, 2015). The most recent MLA report noted a decrease in Chinese enrollments (Looney & Lusin, 2018) at higher education institutions, but research findings from the American Councils for International Education (2017) show healthy enrollments at the pre-college levels and almost all states in the U. S. have programs that offer Chinese. The increasing number of students learning Chinese at the pre-college levels reflects the general recognition of the growing importance of acquiring Chinese language and culture competence for cross-cultural understanding and effective communication in global affairs. The fact that more and more young learners are interested in learning Chinese has led to larger enrollments in advanced Chinese courses, especially in Chinese programs in higher education.

Despite the fact that interest in learning Chinese is steadily growing and the field is becoming increasingly professional, we still need more research that integrates research findings from relevant disciplines such as linguistics, literature, foreign language acquisition, the advancement of educational technology, and techniques that deal with the teaching and learning of advanced Chinese language competency. We have more experience and have done more systematic research on Chinese language pedagogy at the beginning and intermediate levels, but need to pay more attention to teaching methods and techniques for the advanced levels. This article reports on our efforts to enhance our instructional design of the teaching of advanced Chinese though the effective use of currently available technology.

Educational technology has advantages and limitations as well. There is no doubt that the current technology offers us great potential for enhancing the process of teaching and learning of a foreign language (Bai, 2003; Xie, 2001). Multimedia technology can provide learners with multi-stimulus environments because of its capacity of delivering text, graphics, and audio or videos at the same time. Newly developed conferencing programs, such as Zoom, make the offering of distance learning courses much easier. Xie (2001) argues convincingly for Chinese teachers to face the challenge of new technology and to become the "e-generation's Chinese language teachers." He made specific useful suggestions in eight areas where teachers can develop their professional competence in using computer technology to improve their teaching effectiveness. However, most computer application programs or authoring tools are designed for many purposes and often contain more complex features than we need. It requires the collaborative efforts of both IT colleagues and teachers who have demonstrated teaching excellence and know how to integrate technology meaningfully and effectively. Bai (2003) discusses five types of limitations of technology, such as its lack of interactive capacity, its cost, its limited capacity to analyze errors and miscues, etc. He argues that:

> "Only after we fully understand the limitations and advantages of the computer technology can we meaningfully integrate it into our curriculum and help enhance the process of teaching and learning. The key word for successful integration of computer technology into our curriculum is problem-solving, i.e. we start with a pedagogical problem. The specific questions we ask when we think of the use of technology are as follows: 1) Can we do better than what we are doing now if computer technology is integrated? What specific pedagogical problem can it help solve? 2) If the answer is yes to the first question, then we ask the next question. Do we have enough expertise and enough time and money to ensure successful integration of the technology in mind? 3) How can we do it and where it fits into our curriculum" (Bai, 2003).

With generous support from a grant awarded to "The Five Colleges of Ohio and The Ohio State University Postdoctoral Fellowship and Language Enrichment Program" by The Andrew W. Mellon Foundation,[1] we conducted a series of six workshops since April 2017. Our goal has been to identify some of the challenges of advanced Chinese language pedagogy and seek solutions to the problems in order to come up with an enhanced curriculum of advanced Chinese. One of the challenges of teaching advanced Chinese is to meet the diverse needs of students at the advanced levels. We have designed and implemented additional content-based courses to the current curriculum and also explored the possibility of integrating distance learning courses to share among our collaborating colleges. In this article, we report on the instructional design and implementation of an advanced Chinese course in a distance learning format to meet the needs of students who represent a diverse number of academic backgrounds.

## 2. Integrating Distance Learning Technology into the Teaching of Advanced Chinese

As argued in the previous section, we need to start with the identification of a pedagogical problem and then consider the instructional goals, the elements of the instructional design, and how technology can be used effectively to solve the pedagogical problem. The distance learning course we are to report on in this article was designed to solve the problem of 众口难调 [it is difficult to cater to different tastes] at the advanced level. As the number of advanced-level students increases, there is greater demand in upper level Chinese courses for special needs, i.e. advanced Chinese for special purposes. At a small liberal arts college, we cannot offer the variety of advanced Chinese courses on special topics to meet various needs due to smaller enrollments and lack of staffing. In

[1] The authors of the article thank Five Colleges of Ohio, Inc., which sponsored this project with the support of a major grant from "The Five Colleges of Ohio and The Ohio State University Postdoctoral Fellowship and Language Enrichment Program" funded by The Andrew W. Mellon Foundation. Chinese faculty from the OH-5 colleges and The Ohio State University participated in the workshops and contributed to the discussion on how to improve advanced Chinese language pedagogy.

the remaining sections of this article, we will discuss the instructional goals, instructional design, and implementation of a distance learning course for advanced students of Chinese, which enrolls students at two small liberal arts colleges. The immediate benefit of the distance-learning project is that it helps broaden the curricular options at both colleges. We conducted some informal surveys about the effectiveness of the distance-learning course and the reactions are mostly positive.

## 2.1. Instructional Goals

This upper-level Chinese course aims to help advanced students of Chinese develop their advanced language competency through theme-based language learning and independent research projects in Chinese. The students' Chinese proficiency levels of this course range from Intermediate High to Advanced Low, so it was important that we design learning activities according to the students' individual strengths and individual learning goals. The organizing framework that guided our instructional design is the standard-based language teaching and learning framework developed by the American Council on the Teaching of Foreign Languages (ACTFL, 2006). The guiding question we ask is how we can help our students develop their listening, speaking, reading, and writing skills across the three communicative modes, interpretive, interpersonal, and presentational, in the 5 C areas. In the 2012 edition of Performance Descriptors for Language Learners, ACTFL provided a detailed description of those three communicative modes (Table 1).

**Table 1 Three Modes of Communication Proposed by ACTFL[2]**

| Interpersonal | Interpretive | Presentational |
|---|---|---|
| • Active negotiation of meaning among individuals | • Interpretation of what the author, speaker, or producer wants the receiver of the message to understand | • Creation of messages |
| • Participants observe and monitor one another to see how their meanings and intentions are being communicated | • One-way communication with no recourse to the active negotiation of meaning with the writer, speaker, or producer | • One-way communication intended to facilitate interpretation by members of the other culture where no direct opportunity for the active negotiation of meaning between members of the two cultures exists |
| • Adjustments and clarifications are made accordingly | • Interpretation differs from comprehension and translation in that interpretation implies the ability to read (or listen or view) "between the lines," including understanding from within the cultural mindset or perspective | • To ensure the intended audience is successful in its interpretation, the "presenter" needs knowledge of the audience's language and culture |
| • Speaking and listening (conversation); reading and writing (text messages or via social media) | • Reading (websites, stories, articles), listening (speeches, messages, songs), or viewing (video clips) of authentic materials | • Writing (messages, articles, reports), speaking (telling a story, giving a speech, describing a poster), or visually representing (video or PowerPoint) |

[2] Retrieved from ACTFL Performance Descriptors for Language Learners, 2012 edition, https://www.actfl.org/sites/default/files/pdfs/PerformanceDescriptorsLanguageLearners.pdf, accessed Dec.17, 2018.

In the table above, the first row exemplifies the representative performances of each mode, with a more descriptive explanation given in the second row and sample tasks in the fourth row. The third row summarizes specific skills which learners are expected to possess. For example, the typical performance of the interpretive mode is an interpretation of messages. To be more specific, it is one-way communication with no opportunity for negotiation. The sample tasks include reading articles, listening to speeches, and viewing video clips, which are all authentic materials made for native speakers. To accomplish such tasks, the ability to perceive "between the lines" is crucial. It requires learners to assume the perspective of native speakers, which is the main difference between interpretation and comprehension. In the next section, we will explain how the three modes of communication are developed in this distance learning course.

## 2.2. Instructional Design and Implementation

The design of this upper-level Chinese research class aims to assist advanced students of Chinese in developing independent research projects and to help them to reach individual learning goals. Learners go through weekly learning phases to accumulate knowledge and develop language skills necessary for them to complete the research projects and present their findings at the end of the semester. Each weekly instructional phase consists of three major components: 1) an online individual meeting with the instructor; 2) an in-person individual meeting with a teaching assistant, and; 3) an online presentation to their classmates on their work-in-progress, in the form of a symposium, followed by Q&A from their classmates and teacher-led group discussions. Each component requires learners to accomplish different tasks according to individual strengths and academic interests. For example, a learner chose Sino-American relations during Trump's presidential term as her research topic and identified the *New York Times* [Chinese version] as a useful resource. Every week, she selected a relevant news report as the reading material, studied the material before the individual meeting with the instructor, and then prepared for her presentation. She also rehearsed her presentation with a teaching assistant. In addition to uploading her own presentation, she also watched others' presentations before participating in group discussion. In the process, learners practice their skills cumulatively in all three modes of communication—interpretive, presentational, and interpersonal, as illustrated in the table below.

**Table 2 Components of a Weekly Instructional Phase**

| Components | Online individual meeting with the instructor | In-person individual meeting with the teaching assistant | Online presentation followed by a group discussion |
|---|---|---|---|
| Primary Goals | 1. Discussing materials that learners have studied<br>2. Discussing the content and | 1. Correcting mistakes on PowerPoint slides and scripts of presentation<br>2. Rehearsing the presentation and | 1. Recording the presentation with VoiceThread<br>2. Answering questions raised by other fellow students. |

| | | | |
|---|---|---|---|
| | structure of the presentation. | improving pronunciation. | |
| Examples of Tasks (including preparation) | Searching for authentic materials (e.g., news articles, research papers, and videos) related to the research topic; interpreting the main idea of materials; discussing the materials with the instructor. | Preparing PowerPoint slides and scripts of presentation; discussing the presentation with the teaching assistant; rehearsing. | **Presenters:** Recording the presentation; previewing questions raised by the audience; answering the questions; discussing with the audience. **Audience:** Watching the presentations; leaving comments and questions; discussing with the presenters. |
| Communicative modes practiced | Interpretive Interpersonal | Presentational Interpersonal | Presentational Interpretive Interpersonal |

For a distance-learning course, the appropriate use of digital tools is the key to success. We integrated the use of various technological tools, such as VoiceThread, that we mentioned in the table above. The next section focuses on how we incorporated two digital tools, Zoom and VoiceThread, in this distance learning course to bridge the gap between the two campuses.

### 2.3. Implementation of Technological Tools: Zoom and VoiceThread

Zoom is an essential tool for distance learning courses. As a tool designed to provide solutions for online conferencing and webinars, Zoom features many functions which are also convenient for distance learning courses, such as screen sharing, a virtual whiteboard, and Breakout Rooms for simultaneous group work. In the current course, screen sharing and the virtual whiteboard are the two functions primarily utilized. Screen sharing allows participants to view the same webpage or document concurrently, which guarantees that everyone is on the same page. It is also convenient for when the instructor goes through reports or PowerPoint slides with the learners and gives real-time feedback, as illustrated below.

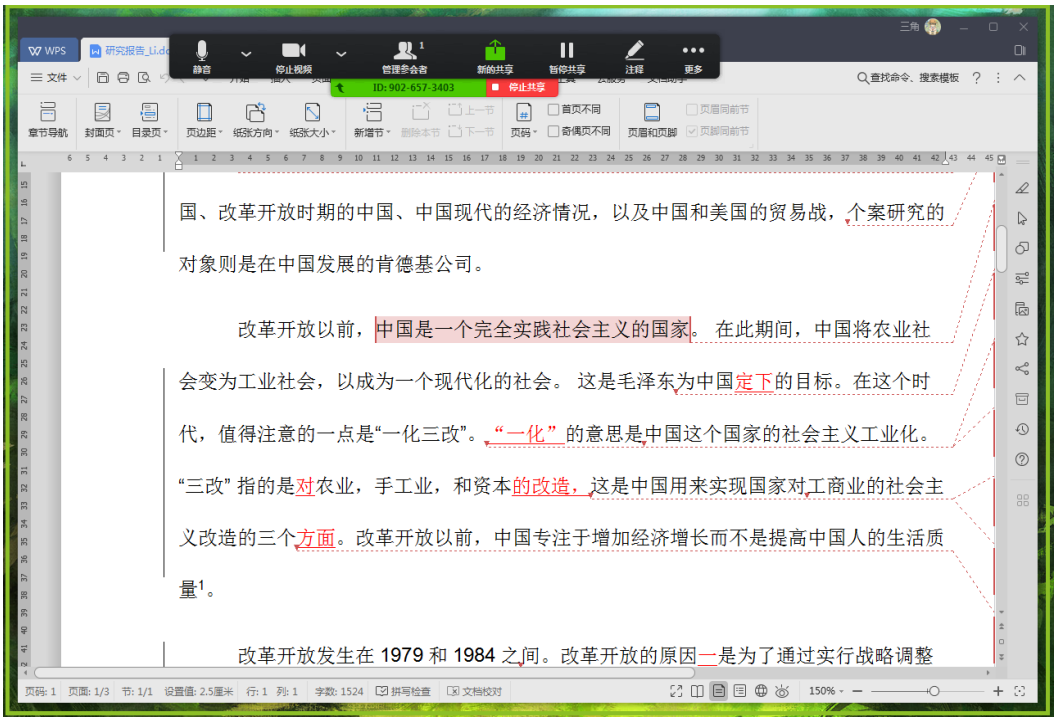



**Figure 1 Using Screen Sharing to Discuss a Learner's Report (the outer green frame indicates the software being shared)**

The virtual whiteboard can be utilized in either individual or group sessions when the instructor invites every participant to contribute to the group class discussion. In the example below, the instructor started a group session by activating the virtual whiteboard and inviting learners to provide some keywords that they remember after watching others' presentations. Participants can either type or draw their answers. They can choose different colors to avoid confusion. With the virtual whiteboard, instructors can design more collaborative activities in a distance learning environment. The following figure is an example of how the virtual whiteboard is used in this distance learning course.

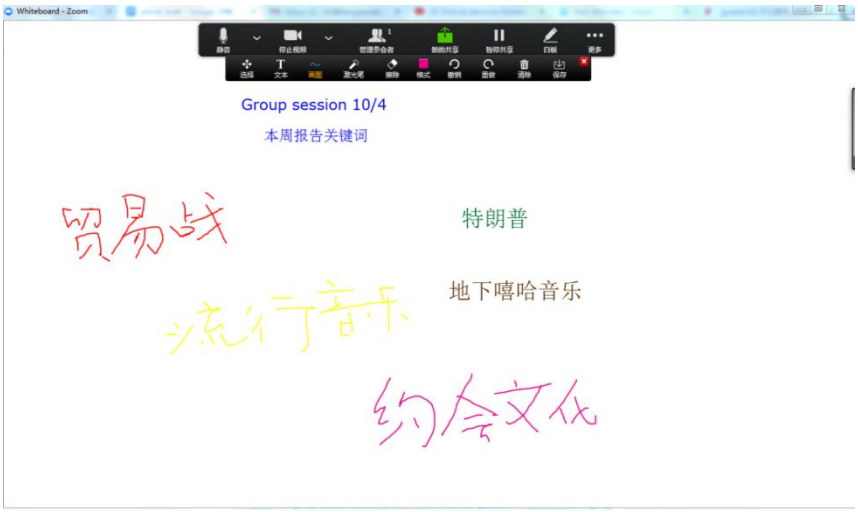



**Figure 2 Using the Virtual Whiteboard to Brainstorm Keywords of Presentations**

VoiceThread is another useful tool we regularly used for the distance learning course. It allows users to record their presentations, either audio or video, and share them with others on the VoiceThread website. Since users can do their recordings slide by slide, it is relatively easy for them to redo a particular section if they feel dissatisfied. As for the audience, they can leave their comments and questions on any slide while listening to the presentation instead of waiting until the end of it. Thus, the presenters can potentially receive more feedback compared to on-the-spot presentations. Researchers in different fields have explored the use of VoiceThread in creating a sense of community in online learning (Delmas, 2017), developing oral proficiency of a second language (Dugartsyrenova & Sardegna, 2017), and promoting collaborative learning (Fox, 2017). However, the use of VoiceThread in learning Chinese as a foreign/second language has yet to be explored. In the current distance learning course, VoiceThread helps ensure high quality, especially the clarity of students' presentations, which can be affected by an unsatisfactory Internet connection. Compared to traditional class discussion, our students spend more time practicing what they wanted to say because they want to show their best on the computer screen for others to view.

Below are two screenshots of a presentation on the trade war between China and the United States recorded by two students. This pre-recorded VoiceThread presentation showcases how the discussion is conducted asynchronously. The research topic chosen by this group of students for this semester is the economic development of China after the implementation of the reform and opening-up policy. For this specific task, they searched for materials related to the recent trade war between China and the United States and made a draft of their presentation for their teachers and classmates to critique. Then, they met with the instructor and teaching assistant to improve the quality of the presentation before they presented their work to the whole class through the use of Zoom. The instructor focused on content and overall structure, while the teaching assistant focused on language-related issues, such as pronunciation and tones or their command of grammar and text structures. After they finished recording their presentation, it was shared among the group, which consists of students from two colleges, the instructor, and teaching assistants. The group members watched their presentation before the group session and left their comments and questions on the screen of their VoiceThread presentation. Thus, the time of group session can be more effectively used as presenters have an opportunity to preview the questions and think about the answers beforehand. Another advantage of using VoiceThread is that the presentations can be easily archived. Learners can review their previous presentations to decide what they can include in their final poster presentation.


**Figure 3 Presenters Attached a Vocabulary List by Adding Comments to a Slide**


**Figure 4 Audience Raised Questions Related to the Current Slide**

## 2.4. Formative and Summative Assessment

The current distance learning course includes an evaluative system that consists of formative assessment and feedback in addition to the conventional summative measures for final course grade. The goal of establishing such a system is providing consistent assistance to the learners as they prepare for their final projects—a poster presentation. The requirement for the final project is introduced to the learners at the beginning of this course, and they are completely aware of the fact that they could not complete the final project and the poster presentation successfully unless they follow the step-by-step instructional guidance and improve their language skills by working through the smaller tasks cumulatively. If we compare this course to a game, the final project would be a formidable enemy who is extremely difficult to defeat, while the learners would be protagonists of the game, who need to gain stronger power to conquer that enemy through accomplishing various quests, such as interpreting the main ideas of authentic

materials and recording presentations. Along the way, the learners might become confused about what to do and how to improve themselves. Therefore, formative feedback is necessary. Learners get formative feedback in both oral and written forms. As mentioned before, learners regularly meet with the instructor and teaching assistant. During those individual sessions, they can receive oral feedback from the teaching team, which helps them elevate the quality of their presentations. They can also gain feedback from other fellow learners regarding which part of their presentations is difficult for the audience to understand. Since the presentations reflect the learning outcome at the end of each learning phase, the learners receive more systematic written feedback after they finish a presentation, including the discussion section in group sessions. The written feedback is meant to provide a comprehensive overview of learners' performance and provide formative feedback to improve students learning. The table below illustrates how we assess students' performance, but it needs to be used more and improved to generate more accurate information as a formative assessment tool.

**Table 3 Sample Rubric for Written Feedback**

| Category | Outstanding | Satisfactory | Decent | Unsatisfactory |
|---|---|---|---|---|
| 内容<br>Content | Demonstrate comprehensive knowledge of the topic. | Demonstrate good understanding of the topic. | Shows good understanding of parts of the topic. | Shows lack of understanding of the topic. |
| 语音、声调<br>Pronunciation and tones | Speaks clearly and correctly almost all (95-100%) the time. | Speaks clearly almost all (90%) the time. Contains errors but with few patterned errors. | Speaks clearly most (85-90%) of the time. Contains some patterned errors. | Low level of comprehensibility due to constant errors of sounds and tones. |
| **词汇、语法**<br>vocabulary and grammar | Able to control structure, vocabulary, cohesive devices with sporadic errors. | Employs vocabulary and grammar structures with occasional errors, but almost no patterned ones. | Employs vocabulary, expressions, and structures with patterned errors that may obscure meaning. | Employs vocabulary, expressions, and structures with many errors that obscure meaning. |
| 理解能力<br>Comprehension | Able to accurately understand and answer almost all questions by classmates about the topic. | Able to accurately answer most questions posed by classmates about the topic. | Able to accurately answer some questions posed by classmates about the topic. | Unable to accurately answer questions posed by classmates about the topic. |

Figure 5 below is an example of written feedback that learners receive after a presentation. Instead of points (e.g., 100, 90, 80, etc.) or grades (e.g., A, B, C, etc.), learners get a rating for each category listed in Table 3. For example, this learner received "outstanding" for pronunciation because only a few mistakes were made throughout the whole presentation. Comprehension is also "outstanding" because the discussion part went smoothly, with the learner properly elaborating on her points. However, she received "satisfactory" for content because her presentation did not reflect a comprehensive understanding of the subject matter of her research. She also made some statements with insufficient explanations. The primary purpose of the formative written feedback is to raise learners' awareness of their strengths and weaknesses in each aspect of their presentation, such as pronunciation, structure, and content, that should be improved in their future presentations.

The special task is another crucial feature of this evaluative system. It is designed to push learners to fix their urgent problems because they must finish their tasks and report the results back to the instructor to claim the "completion" of a learning phase. The tasks are assigned based on each learner's specific performance during that week and listed at the end of written feedback. Most tasks aim to help learners improve the quality of their presentations. For example, in Figure 5, the learner's special task is adding a slide according to the instructor's feedback as a final revision of her presentation. To finish this task, the learner has to revisit certain slides and arrange them in a more reasonable way. The task can also raise her awareness of similar issues in the future. The learners may also receive other types of tasks, which aim to regulate their learning behaviors. For example, some learners are required to leave at least one comment on each of the others' presentations, usually because they forget to do so during a particular week. The special task is an essential component of this formative evaluative system because it creates a situation in which the learners must fix their existing problems before proceeding to the next learning phase. Moreover, it helps learners develop better learning behaviors and reduces the number of recurring mistakes.

10.18 Presentation

Pronunciations and Tones: Outstanding
Slide 4: 躲 ( duo3 ) 避 ( bi4 ) ; 自 ( zi4 ) 我; 低 ( di1 ) 俗
Slide 5: 掌 ( zhang3 ) 控 ( kong4 )

Content: Satisfactory
Good content and structure. I have two suggestions:
1. You mentioned three ways to avoid censorship, but I was not sure whether "自我审查" counts as one way to do that. It would be clearer if you have a slide entitled "回避政府审查的方式", on which you can list all three ways, and then proceed to explain each of them with three separate slides.
2. In both your introduction and conclusion part, you mentioned that your main topic is "政府利用嘻哈音乐获得支持". However, you spent half of your presentation to explain how to avoid censorship. That is very confusing.

Vocabulary and Grammar: Satisfactory
Slide 1: "什么的" is too casual for a formal oral presentation. You can use "之类的"
Slide 2: "马克思是对的" → 《马克思是对的》; "一带一路"的节目 → "一带一路"的战略; 中国政府的传统美德 → 中国政府倡导的传统美德; 习近平总统 → 习近平总理
Slide 3: perspective: 角度
Slide 4: 躲避政府审查 → 避免受到政府审查

Comprehension: Outstanding

Overall: Satisfactory

Your special task for next week:
Per my first suggestion under "content", please add a slide before/after that on 自我审查.

**Figure 5 An Example of Completed Formative Feedback Form**

## 3. Integrating Technological Tools in the TA sessions

The TA session is a face-to-face, one-on-one discussion after students complete their on-line lecture class with the course instructor. The in-person individual meetings aim to help students improve their Chinese language control, such as their pronunciation and tones, helps foster correct use of words, grammar, and text structure, and gets students to focus on the correct usage of language at the advanced level. In addition, the one-on-one sessions also help students develop their independent learning strategies. We design and implement tasks to increase the interaction between the TA and students. The TA is a facilitator of effective learning rather than only a provider of input or a grammar

checking machine. During the 30-minute class, the teaching assistant helps students polish their writing by giving students corrective feedback on their presentations and providing them with online tools, such as BCC corpus and Sketch Engine[3] corpus, to facilitate students' learning of strategies on how to improve their advanced Chinese competency effectively.

### 3.1 Incorporating the Use of a Corpus in TA Sessions

Several scholars have advocated using corpora to aid language learning (Bernardini, 2004; Leech, 1997; McEnery, Xiao, & Tono, 2006; Yeh & Zhang, 2018). By using a corpus, a student can look up a word's precise denotation and connotation through the use of tools such as concordances, and compare its usage in different linguistic or pragmatic contexts (Leech, 1997, p. 8).

Yeh and Zhang's study (2018) suggests that the use of a corpus helped students improve their storytelling abilities. The use of a corpus allows the student to understand semantic and pragmatic features, and also grammatical characteristics, such as the common collocating words based on the context in which the word is used. For instance, when the two students were preparing their presentation on the US-China trade war, they learned how to use the corpus tool to improve their successful choice of words to describe the wave of tariffs by searching "Verb+tariff" in BCC corpus to find the appropriate collocating verbs. In Figure 1, the corpus shows a list of examples of using "tariff" as an object, such as "canceling the tariff" [取消关税], "levying the tariff" [征收关税], and "lowering the tariff" [减少关税]."


**Figure 6 Inquiry *Verb+tariff* in BCC corpus**

Clearly, the corpus is a great tool for students to independently improve their vocabulary control and learn the contextual forces upon word choice. However, students

[3] c.f. http://www.sketchengine.eu

often still need a TA's guidance to learn how to extract from the concordance and discover words from the corpus' data. This is why Leech (1997) stresses the importance of teaching students the ability to use or "exploit" corpora. McEnery, Xiao, and Tono (2006) continue this line of thought by saying that knowing how to use a corpus enables students to independently develop their own studies or research. Once students master the necessary knowledge and skills, language study "may become student centered" (p. 83).

Bernardini (2004) focuses on corpus-aided discovery learning. In corpus-aided discovery learning, teachers do not present the rules to students, but guide students to explore a range of examples from corpora so as to discover the characteristics of a linguistic phenomenon. As mentioned earlier, Yeh and Zhang (2018) successfully apply this method to improve student's storytelling abilities. Their corpus-based instruction on discourse-linking *jiu* (就) consisted of four 50-minute classes. Each teaching unit included five steps: warm-up and review, awareness-raising, analysis, production, and assessment.

While discovery learning can be valuable in our independent study class, improving students' problem-solving and autonomous learning skills are especially important in TA sessions. Based on the goals of the independent study session, the ability to improve independent study and the role of the TA-sessions (i.e. to solve problems), and considering Yeh and Zhang's five step method, we formulated the following instructions to support successful "exploitation" of these technological tools.

In this case, we first noted the students' difficulty in finding the exact expressions to talk about the current tariff situation between the United States and China. Our TA designed learning tasks to help our students develop the ability to use the corpus tool to explore the semantic, grammatical, and pragmatic feathers of the key word "tariff [关税]."

During the one-on-one or group discussion sessions, our TA provide step-by-step guide to help students develop the ability to look for the answers on their own. For example, when students tried to describe the tariff they produced sentences like "have lower tariffs" [有更低的关税], "President Trump makes the tariffs higher than before" [特朗普让关税比以前高]. They use grammatically correct sentences but inappropriate expressions to express their intended meaning. In order to improve their word usage, the teacher shows students how to look for the word in a corpus and discuss how to retrieve the useful information from the mass language data bank that the corpus provides.

Another example was to help an advanced student of Chinese with his independent research paper. He tried to describe the development of Rock-n-Roll music in Mainland China and Taiwan. However, in the process of describing the characteristics of the music, the student did not have adequate vocabulary, and therefore repeatedly said "*qingsong de yinyue"* 轻松的音乐 [light music] and "*bu qingsong de yinyue"* 不轻松的音乐[not light music]. In the first chapter of the book *Teaching Collocations*, Morgan Lewis points out that in order to turn students into advanced learners, it is a teacher's job to point out more precise or concise options to improve a student's performance (Lewis,

2000, p. 22). To enhance precision, Lewis' preferred method is collocation, because, as elaborated in chapter three by Jimmie Hill, "the way words combine in collocations is fundamental to all language use" (Hill, 2000, p. 53). In this case, using Sketch Engine to look up collocates of the word "*yinyue"* 音乐[music] can give the student a wide range of adjectives describing music's timbre, such as *youyang* 悠扬[melodious] and *huankuai* 欢快 [cheerful].


**Figure 7 Inquiry of adjective modifiers of *music* in Sketch Engine**

### 3.2 Learning Semantic, Pragmatic and Stylistic Features from Context

It is important to make students aware of the influences that contextual forces exert on the usage of words. Even though a word or a sentence might be grammatically correct, it might be inappropriate pragmatically due to lack of in-depth understanding of the word or sentence structure. When facing pairs or groups of words that are seemingly synonymous, students may have difficulty finding the exact word to express their intended meaning. In this case, the teacher can help students learn how the seemingly synonymous words differ through the use of a corpus, which helps to differentiate similar words by showing the actual collocations. For example, a student did a presentation on the development of the Chinese economy and wrote the following incorrect sentence

* 国家*创造*了委员会，进行经济体制改革的总体设计。

to mean "the country established a council, and started the design of the economic system reforms." In this sentence the correct verb for "to establish" would be *chengli* 成立 [to establish, to found) and not *chuangzao* 创造 [to create]. Our TA guided the students to compare these two words by showing them the usage difference in Sketch Engine. In the screenshot below (see Figure 8), the left-side collocates are more closely related to *chengli*, such as *xiaozu* 小组 [group]s and *weiyuanhui* 委员会 [committee, council]; the right-side collocates to *chuangzao*, such as *tiaojian* 条件 [condition] and *jihui* 机会 [opportunity]. After observation and discussion guided by the teacher, students can infer

the rule that when talking about establishing an organization, it is correct and appropriate to use *chengli* (成立).

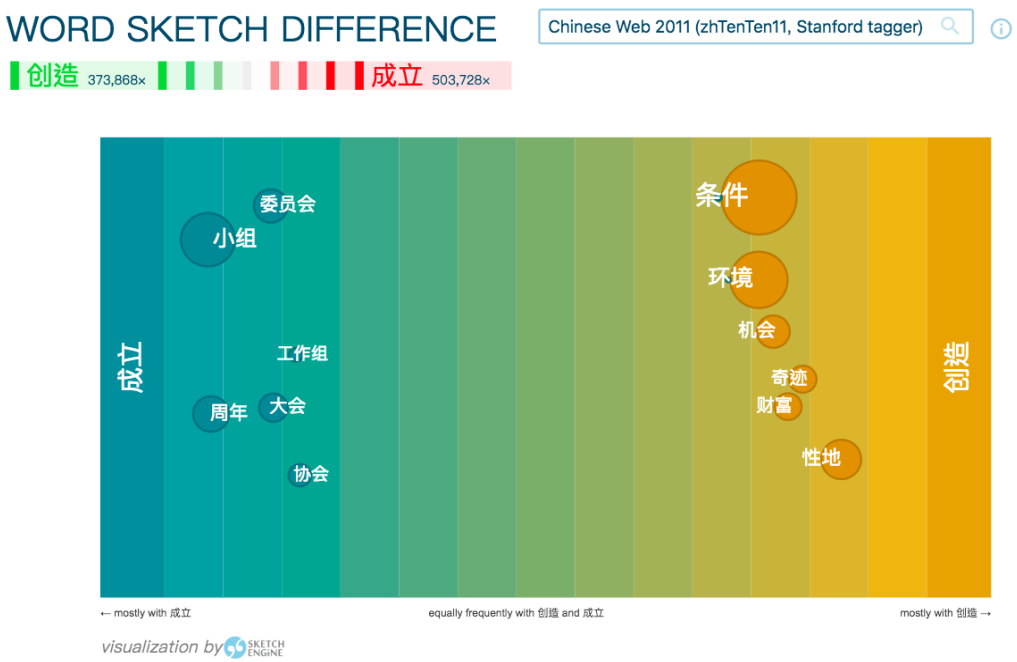



**Figure 8 Comparing and Contrasting "chengli, chuangzao" Visually in Sketch Engine**

### 3.3 Corpora as a Tool in Effective Lesson Planning

A corpus is a handy tool for teachers when they prepare their lesson plans. For instance, it allows teachers to extract authentic examples of the usage of a lexical item, which provides a more objective view of language use than that of pure intuition or introspection. In addition, the corpus can also retrieve good dictionary examples and help teachers select the optimal usage examples. Teachers can also get statistical information from corpora that provide informative evidence for accurate word choice. It can successfully help teachers be more confident in advising students on word usage because of the informative evidence on the word's semantic features, contextual forces, and Chinese's regional variations.


**Figure 9 Authentic examples of "Trade War" on the Chinese Web 2011, Sketch Engine**

## 4. Conclusion

This article is a report on our efforts to enhance our Chinese language curriculum by designing and implementing a distance learning course that meets the diverse needs of students of advanced Chinese. This upper-level Chinese course aims to help advanced students of Chinese develop their advanced language skills across the 3 communicative modes, interpretive, interpersonal and presentational, and develop their analytical, critical and creative thinking skills through project-based learning. Based on our regular reflections among the teaching team members and our conversations with the students at the end of the semester, we feel very positive about the addition of this new course to our Chinese language curriculum. We have decided to make this course a permanent element of the advanced Chinese curriculum and plan to explore more opportunities to work collaboratively with institutions beyond the OH-5 colleges in the future.

## References

American Councils for International Education. (2017). The national K-16 foreign language enrollment survey report. Retrieved from https://www.americancouncils.org/sites/default/files/FLE-report-June17.pdf

Bai, J. (2003). Making multimedia an integral part of curricular innovation. *Journal of Chinese Language Teachers Association, 38*(2), 1-16.

Bai, J. (2007). Overview: Past, present, and future [of teaching Chinese as a foreign language]. In M. Tseng (Ed.), *AP Chinese language and culture teacher's guide* (pp. 1-7). New York: College Board.

Bernardini, S. (2004). Corpora in the classroom: An overview and some reflections on future developments. In J. Sinclair (Ed), *How to use corpora in language teaching* (pp. 15-36). Amsterdam: John Benjamins Publishing Company.

Delmas, P. M. (2017). Using VoiceThread to create community in online learning. *Techtrends: Linking Research and Practice to Improve Learning, 61*(6), 595-602.

Dugartsyrenova, V. A., & Sardegna, V. G. (2017). Developing oral proficiency with VoiceThread: Learners' strategic uses and views. *Recall, 29*(1), 59-79.

Fox, O. H. (2017). Using VoiceThread to promote collaborative learning in on-line clinical nurse leader courses. *Journal of Professional Nursing, 33*(1), 20-26.

Goldberg, D., Looney, D., & Lusin, N. (2015). *Enrollments in languages other than English in United States institutions of higher education, Fall 2013*. Retrieved from https://apps.mla.org/pdf/2013_enrollment_survey.pdf

Hill, J. (2000). Revising priorities: from grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: further developments in the lexical approach* (pp. 47-69). Boston: Thomson Learning, Inc.

Looney D., & N. Lusin. (2018). *Enrollments in languages other than English in United States institutions of higher education, Summer 2016 and Fall 2016*, Preliminary Report. Retrieved from https://www.mla.org/content/download/83540/2197676/2016-Enrollments-Short-Report.pdf

Leech, G. (1997). Teaching and language corpora: A convergence. In A. Wichmann, S. Fligelstone, T. McEnery, & G. Knowles (Eds.), *Teaching and Language Corpora* (pp. 1-23). London: Longman.

Lewis, M. (2000). There is nothing as practical as good theory. In M. Lewis (Ed.), *Teaching collocation: Further developments in the lexical approach* (pp. 10-27). Boston: Thomson Learning, Inc.

Ling, V. (2018). *The field of Chinese language education in the U.S.: A retrospective of the 20th century.* London and New York: Routledge.

McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book.* London: Routledge.

National Standards in Foreign Language Education Project. (2006). *Standards for foreign language learning in the 21st century (3rd Ed.).* Yonkers, NY: American Council on the Teaching of Foreign Languages.

Xie, T. W. (2001). E-generation's Chinese language teachers: Meet the new challenges. *Journal of the Chinese Language Teachers Association, 36*(3), 75-80.

Yao, T. C. (1996). A review of some computer-assisted language learning (CALL) materials for Chinese. In S. McGinnis (Ed.) *Chinese pedagogy: An emerging field* (Chinese Language Teachers Association Monograph Series #2) (pp. 255-284). Columbus: National Foreign Language Resource Center at the Ohio State University.

Yeh, M., & Zhang, X. (2018). Corpus-based instruction: Teaching discourse-linking *jiu* (就) in storytelling. *Chinese as a Second Language*, *53*(1), 1-24.

# 中文线上课堂有效结合科技工具以强化互动之报告
# (Enhancing Interaction through the Effective Incorporation of Technology Tools for a Virtual Chinese Language Classroom)

曾妙芬
(Tseng, Miao-fen)
弗吉尼亚大学
(University of Virginia)
mt3z@virginia.edu

高燕
(Gao, Yan)
亨利科县公立学校
(Henrico County Public Schools)
ygao@henrico.k12.va.us

蔡罗一
(Cai, Luoyi)
北卡罗莱纳大学教堂山分校
(University of North Carolina, Chapel Hill)
luoyicai@email.unc.edu

**摘要：**本篇论文旨在报告 2018 年美国弗吉尼亚大学星谈中文项目中线上科技辅助工具的有效结合与应用及其对网络中文课堂互动性的促进作用。弗吉尼亚星谈项目于 2018 年首次在网络中文课堂的教学过程中，有机结合在线视频工具 ZOOM 和教学互动软件 Nearpod 的各项功能，例如标注、即时测试、在线同步合作等，针对具体语言教学目标设计课堂活动，优化了学生在课堂中的语言输出，增强了学生学习的效率和趣味性，弥补了线上教学互动性不足的缺陷。根据项目的后期反馈，科技工具的应用有效地促进了学生的学习效果，提高了学生对课程的满意度。

**Abstract**: This paper reports on the effective incorporation of online technology tools in the University of Virginia's STARTALK Chinese Student Program. These tools contributed most importantly to the interactive component of online Chinese language classrooms, which can otherwise be absent or minimal. In 2018 the Virginia STARTALK program created a streamlined online learning process, organically strengthened by the combination of the online platform ZOOM and the teaching interactive software Nearpod, whose features include annotation, real-time testing, and online synchronization cooperation. These classroom activities were designed to fulfill specific learning objectives, achieve optimal language output, and enhance student learning efficiency and engagement. The application of technology tools has effectively strengthened online interactions between instructor and learners, thereby promoting more effective language learning and improving student satisfaction with the course.

**关键字：**线上教学、科技辅助工具、线上课堂活动设计、线上互动、星谈项目

**Keywords**: Online teaching and learning, technology tools, online activity design, online interaction, STARTALK

## 1. 前言

近年来，科技辅助工具在语言教学领域的应用越发广泛，围绕科技辅助工具在语言教学中的应用这一话题，学界讨论的重点也早已从“是否要将其应用于语言教学”转向“如何将其应用于语言教学”以增强学生学习效率（Warschauer & Meskill, 2000; Hoopingarner, 2009）。除了如何应用科技辅助工具之外，在线上语言教学领域，普遍存在的另一个重要问题在于课堂的互动性不足。在线上教学场景中，因为时间、空间和信息技术等方面的限制，教师与学生、学生与学生的沟通效率不如线下面对面教学，因此，如何通过科技辅助工具增强沟通效率，提高课堂互动性，最大化学生在课堂中的语言输出以提高学生的语言学习成果已成为学界和各语言项目讨论和研究的重点。

本论文结合 2018 年星谈中文项目的教学实践，旨在讨论如何通过多种科技辅助工具结合的方式来设计针对具体语言目标的课堂活动，以增强网络语言课堂的互动效率和趣味性、优化学生即时交流中的语言输出质量。

## 2. 星谈项目背景

弗吉尼亚大学星谈项目始于 2008 年，每年均获得美国联邦政府经费，开设针对 9 至 12 年级中学生的中文课程，学生项目总计 12 天。自 2016 年起，该项目不授予任何学分，完全采用线上教学的模式，所招学生遍布美国各州。学生在进入项目之前需完成高中中文二、中文三或同等课程。根据美国外语教学学会外语水平大纲要求，项目前学生的水平在初级高到中级初之间，项目结束后，希望达到中级初水平，迈入中级中水平。除招收学生外，项目也在全美范围内招收 12 名中文教师，这些老师既有从事 K-12 阶段教学的，也有在北美大学中文项目任教的。在学生项目正式开始之前，由项目主管老师团队对所招收的 12 名教师进行为期三周的线上线下相结合的培训，1 名教师在通过培训之后进入学生项目，对学生进行实际的线上教学。

2018 年弗吉尼亚大学星谈学生项目共招收了 24 名学生和 12 名老师。根据 24 名学生的学习背景和语言水平，两名学生为一组分成 12 组；教师团队则分成相应的 6 组，每组两位老师；每位老师轮流对所分配的两名组内学生进行每天一小时的线上教学。依据翻转课堂的理念，项目要求学生在每次上课之前和之后完成相应的课前及课后语言任务。与此同时，应差异化教学所需，项目也为学生开设一对一的线上课外辅导，由学生自主决定是否参加并提前预约辅导时间。

## 3. 星谈项目科技辅助工具概览

2018 年星谈项目主要采用了包括 Canvas、Quizlet、ZOOM、Nearpod 和 Padlet 等 8 种不同的科技辅助工具，贯穿学生课前、课中到课后三个阶段的学习。表 1 简要概括介绍了 8 种科技辅助工具在不同阶段的应用情况：

**表 1 课前—中—后科技辅助工具的应用和功能**

| 课程阶段 | 科技辅助工具 | 主要功能 |
|---|---|---|
| 项目期间 | Canvas | 学习资源中心，提供课程大纲、每日学习计划、课件和教师反馈 |
| 课前 | Quizlet | 提供课前预习任务的词卡和词汇自测练习 |
| | PlayPosit | 提供课文导读、讲解视频等学习资源 |
| 课中 | ZOOM | 主要作为网络课堂的平台，提供标注、文字输入、即时问答等课堂活动辅助功能 |
| | Nearpod | 支持多种形式的网络课堂活动，增强网络课堂互动性 |
| 课后 | Padlet | 学生完成并提交课后写作任务的平台 |
| | Flipgrid | 学生完成并提交课后口语任务的平台 |
| | Facebook | 师生、生生课后互动社区 |

文章下一部分将对 ZOOM 和 Nearpod 做详细介绍，此部分简要介绍 Canvas 和 Quizlet 等其他六个工具及其在星谈项目中的运用情况

### 3.1 Canvas[1]

Canvas 是美国 Instructure 公司的一个学习管理系统，和 Moodle，Sakai 一样是一个开源的学习管理系统。在 Cavans 主页注册免费帐号，通过直观的用户界面对课程教学进行管理操作，Canvas 可以用于在线，混合或面对面教室中。与其他的学习管理系统相比，Canvas 完全架构于云服务之上，不受硬件环境约束；对移动应用、互联网应用有更深度的整合。Canvas 可提供全方位的服务，包括云管理、存储与共享、创建与编辑文件，还可以通过云端的 Canvas APP 中心嵌入各种在线

[1] c.f., https://canvas.instructure.com/login/canvas

应用以达到翻转教学的目的。本项目进行期间，学生通过 Canvas 学习管理系统了解当日学习计划，Canvas 除了与其他学习管理系统一样为用户提供一整套在线课程管理的基本功能，如课程作业、测试，小组讨论，学习评价与课程内容工具，学习支持工具等，还拥有 App 中心，支持大量工具嵌入使用，用简单直观的，用户喜欢的方法提供更加流畅的在线管理功能，如提供成绩单，评估工具，讨论工具，在线聊天，视频交流等功能，并支持拖拽文件上传，此外 Canvas 用户还可以通过 Facebook，SMS 文本信息，email 和其他通讯渠道获取 Canvas 通知。

图 1 中显示的是 Canvas 课程主页中的课程导航、内容范围和侧栏。课程导航包括帮助学生前往课程具体地方的链接。主页内容（和所有 Canvas 内容）显示在内容范围区域里。内容可以按照个人意愿安排课程布局的设置，包括页面、教学大纲、讨论、公告、测验或已导入的内容。主页也可在页面顶部设置显示最近的公告。主页内容还指明哪些部分会在侧栏显示。侧栏与控制面板 (Dashboard) 侧栏的功能一样，但只显示特定课程的内容并包含额外选项。侧栏的顶部包含管理课程的工具。侧栏的底部包含待办事项的清单，其它部分则取决于课程主页设置的布局。要使用测试学生浏览 Canvas 课程，可以点击学生视图 (Student View)链接。一旦发布了课程，并且学生已经开始参加课程，则可以利用查看课程分析 (View Course Analytics) 链接来访问课程的分析结果。侧栏显示待办事项部分，并显示 Canvas 中所有需要评分的项目和截止日期。

Canvas 与其他常见的学习管理系统相比，最大的优势在于可以有序、清晰地囊括全方位的服务，包括云管理、存储与共享、创建与编辑文件等。本次星谈项目在云端的 Canvas APP 中心嵌入各种线上应用辅助工具：如 Quizlet、Playposit、Padlet 和 Fligprid，以翻转教学的理念，通过丰富多样的形式设计课前及课后的语言任务。项目进行期间，学生只需登陆 Canvas 学习管理系统，即可一目了然地了解当日学习计划，每周课程安排，获取相关学习资料，并根据要求点开罗列好的相关链接，使用多种线上辅助工具完成覆盖听、说、读、写的课前及课后语言任务。

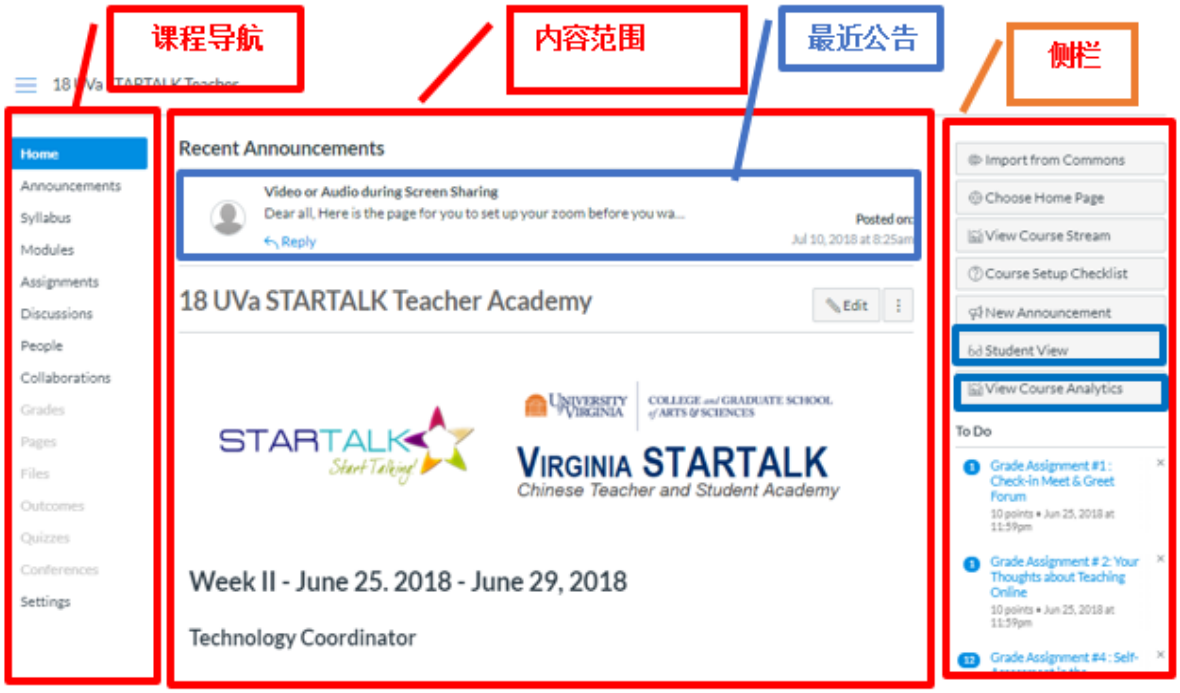



**图 1 Canvas 课程主页**

### 3.2 Quizlet[2]

Quizlet 的创办人，Andrew Sutherland，最初是为了准备高中的法语单字创造了这个学习工具。如今 Quizlet 已经成为全球最大的教师与学生线上学习社群，主要功能就是为大家提供词卡学习语言。任何人都能用 Quizlet 通过“搜索”功能来学习其他用户所创建的内容，或者通过“建立”功能来创建自己的词汇库，也可以与其他人分享所学习的内容。建立学习集时，免费 Quizlet 的用戶可以从“我们的图库”添加图片；升级过的用户除了可以使用“我们的图库”，还能上传自己的图片建立自己的学习集。为了管理方便，老师可以创建班级或者文件夹，所创建的班级或者文件夹可以生成一个分享链接与班级的同学分享，任何人收到该 URL 即可注册并自动加入你的班级。此软件还秉承寓教于乐的理念通过趣味样的让学生自主选择学习词汇的方式如拼写、测试、配对等（图 2）。本次星谈项目中，学生们的课前预习任务中就包括使用 Quizlet 预习与第二天学习主题相关的重要生词并进行自测。学生可根据自己的学习方式选择不同的自测模式，老师和项目助理则可通过教师版 Quizlet 追踪学生的学习情况。

Quizlet 除了运用在课前预习任务的设计上，也可运用在网络中文课堂中。本次星谈项目的线上教学部分运用了 Quizlet 最吸引人的 Quizlet Live 功能。在最后的总结回顾课上，教师将参与课程的所有学生聚到同一个 ZOOM 教室，通过 Quizlet Live 的随机组队功能让学生们组队进行竞赛。比赛过程中，学生需要将随机抽取的

[2] c.f. https://quizlet.com

重点词汇和定义进行正确配对，第一个正确匹配完所有词汇和定义的团队赢得比赛，如果团队中有一人答错则这个团队的成绩进度条退回原点，需要重新迎头赶上。随机组队的功能让学生们能够与日常小班之外的学生进行互动，且小组内的每一位成员都需要为赢得比赛作贡献，随时可以看到的进度条和背景音乐增加了比赛的紧张感和现场感，让每一位学生时刻保持专注。通过 Quizlet Live 设计课前热身、课后复习的词汇比赛，不仅让学生在趣味十足的组队竞赛中不断复习学习重点、快速进入状态，营造出的激烈紧张感，也使得教师更容易掌握课堂节奏，抓住每一位学生的注意力。

另外，付费升级版的 Quizlet 还可以穿插更多富有吸引力的内容，创建自定义团队，玩图标游戏，添加音频等，对激发学生的自主学习兴趣、活跃课堂气氛起着积极作用。

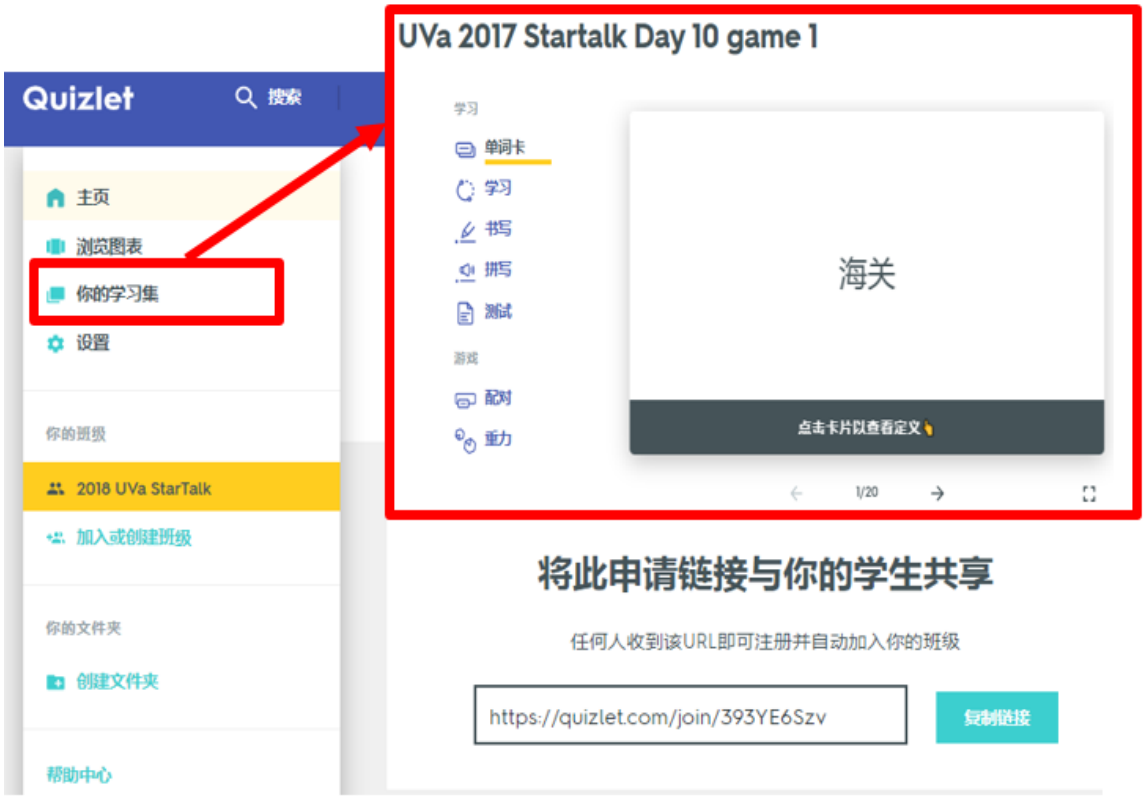



**图 2 Quizlet 操作界面**

### 3.3 Playposit[3]

PlayPosit 以前被称为 eduCanon，是一个在线的课堂教育视频分享平台，有免费版和付费版。目前 PlayPosit 可以通过多种形式选择视频网站的视频内容（图 3），

[3] c.f. https://go.playposit.com/

运用已有的视频，创建交互式视频用于教学。在构建交互式视频过程中，可以向视频添加多种形式的问题，因而学生可以在视频推进过程中参与互动问答（图 4），从而促进翻转课堂的有效实施，使学习者积极主动地参与到学习中。PlayPosit 在构建视频过程中可选择多个视频，也可以添加已有模板。在插入互动问题时也有多种选择，既可以灵活地控制互动问题的切换时间也可以设置持续时间。新增加的标签功能可以为做好的视频活动添加标签以便以后查找方便。PlayPosit 作为教学工具的另一有效功能就是老师可以及时看到学生答题情况，提供必要的可操作数据，以便制定下一步/课程计划。本次星谈项目通过 Playposit 的视频互动问答为学生们设计了每日课前导读、语法讲解视频等一系列有针对性的课前预习资源，以帮助学生更细致、全面地为第二天的线上课程作准备；与此同时，教师在第二天教课前也可通过 PlayPosit 平台快速收集、了解学生的答题情况，并有的放矢地对线上教学内容进行及时调整和补充，根据学生的答题情况在线上教学中给出更为有效的反馈。

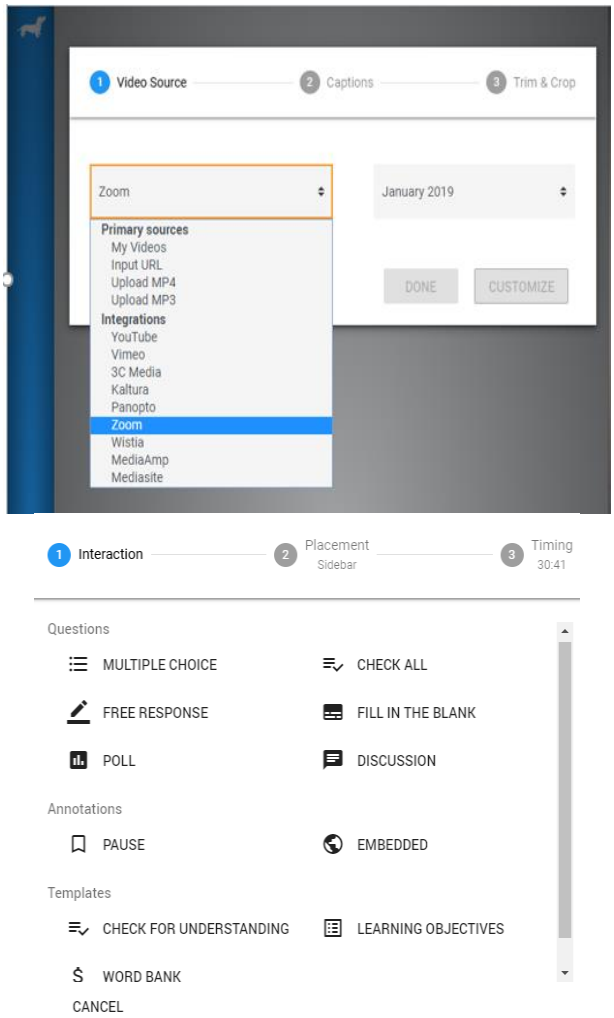



**图 3 Playposit 创建互动视频界面**


**图 4 Playposit 学生回答互动问题**

### 3.4 Padlet[4]

Padlet 是一个在线协作工具，只需要通过简单的操作就可以做出一个 “故事墙” 。不过目前 Padlet 免费版只可以创建三个 “故事墙” ，如果要建立更多的，就需要选择付费版本。该工具包含丰富的模板，支持多人实时协作，支持多种形式展示（图 5），很适合作为教学时的互动小工具。此外，Padlet 设有网页版（网址：https://padlet.com/）和手机版两种版本。建立好了 “墙” 之后，只需双击就可以在 “墙” 上编辑自己的内容，点击“分享”按钮，就可以把 “墙” 分享出去，所有得到链接的人就可以在 “墙” 上进行互动。鉴于 Padlet 易于分享协作、展示成果的功能，本次星谈项目将 Padlet 运用在课后写作任务的设计中。根据每日线上教学的主题学生以短博客、书信等形式写一段话，如图 5，学生以“跟中国朋友去 KTV 唱卡拉 OK 的经历”为主题，分享了各自的“经历”。Padlet 直观的界面，让每一位学生在完成自己的写作任务的同时，也能看到其他同学的作品，并相互点赞，增强了师生、生生之间在课外的互动。

[4] c.f. https://padlet.com/

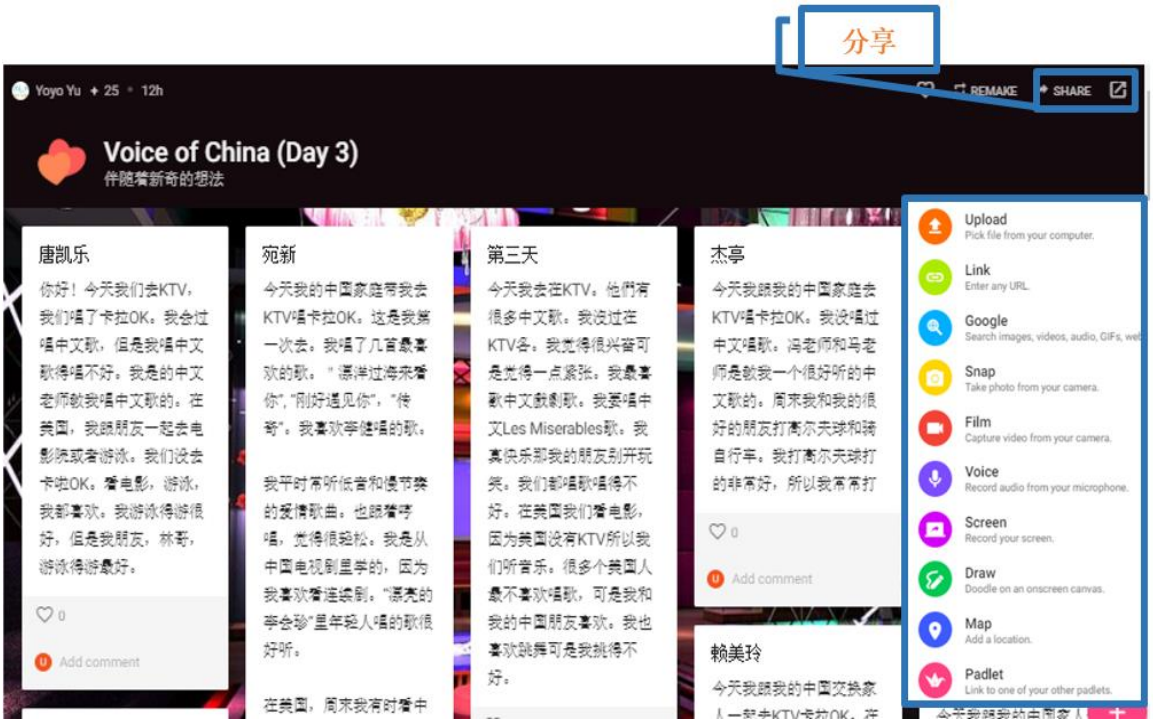



**图 5 Padlet “故事墙”**

### 3.5 Fligpgrid[5]

Flipgrid 于 2015 年成立于美国明尼安波利斯，由明尼苏达大学的教授和研究生创建。目前 Flipgrid 的用户可以完全免费使用服务。Flipgrid 是一个让教师通过建立数字社区，以短视频为工具，让学生进行社交学习的应用。每一位老师都可以在 Flipgrid 建立自己的教室，并在教室里发布讨论主题，学生可以通过制作短视频来分享自己的观点，视频内容可以被教室里的所有用户看到。每个教室都有密码确保安全性，老师可以将密码或者链接分享给自己的学生。学生虽然可以随意上传视频，但是老师有权对学生的视频进行审核、编辑或者删除，从而有效地控制上传内容的质量。与 Padlet 类似的优点是，Flipgrid 可以由老师选择开放视频的点赞、评论功能，这以此激发学生的积极性。与使用 Padlet 布置课后写作任务相对应，本次星谈项目选择使用 Flipgrid 设计课后口语任务。学生在制作视频的过程中，不仅锻炼了自己的口语表达能力，而且充分发挥了的创造力和表现力，学生们制作的视频有一人分饰两角的原创小品，也有自弹自唱将英文歌翻译成中文的，还有很多精心制作给视频加入后期特效的，形式丰富多样，内容精彩纷呈（图 6）。

[5] c.f https://flipgrid.com/

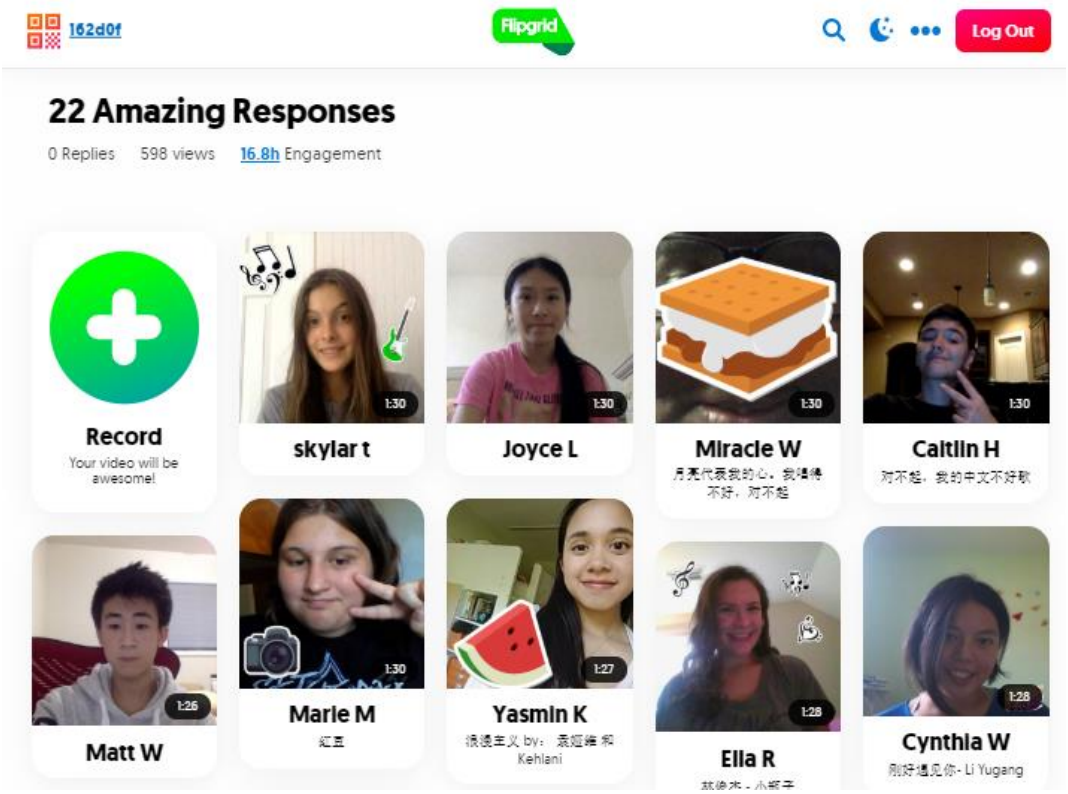



**图 6 Flipgrid 学生分享视频**

### 3.6 Facebook[6]

Facebook作为广为人知的社交媒体平台，近年来受到越来越多语言项目的青睐。相较于在海外留学项目被更多运用的微信（WeChat），Facebook 平台在美国学生中的普及率更高，信息分享也更公开，因此更适用于面向美国本土学生的线上语言教学项目，它能够为实际“相隔千里”的学生们在课堂之外提供一个“近在咫尺”的交流空间，在这个空间内，学生们可以在课堂之外接触到更丰富的真实语料，在语言学习中了解文化，并且与整个项目中来自全美各地的学生交流、分享他们的看法。基于上述优点，本次星谈项目利用 Facebook 平台开辟了“2018 UVa STARTALK Chinese Student Academy”线上小组，要求全体学生、项目管理团队和任教老师都加入该小组并参与到每天课后的话题互动中。每天的线上同步课程结束后，学生除了在其他平台（Padlet 和 Flipgrid）完成书面和口头的语言任务之外，还需要去Facebook小组浏览当天的小组话题。星谈项目的学生助理会根据每天网络课堂的教学主题选取相关的真实语料发起一个讨论话题，如图 7：

[6] c.f. https://www.facebook.com/


**图 7 Facebook 小组课外互动话题**

学生需要在项目助理发布的与课程相关的讨论话题中选取至少 6 个有兴趣的话题进行文字回复，同时还要回应其他同学分享的看法，可以简单的点赞，也可以文字评论。学生的每个回复都会由项目助理阅览和随机回复，并将值得注意的重要信息反馈给相应的任课老师。同时，进行线上教学的任课老师也可以随时去 Facebook 讨论小组浏览学生们的回答，并在需要的时候即时在 Facebook 平台给出

评论和反馈。在本次星谈项目中，Facebook 小组功能的运动无疑对扩展和丰富课堂教学内容，增进师生、生生在课堂外的了解与互动起到了积极的推进作用。

在本次星谈项目的线上部分，教师则以在线视频工具 ZOOM 为平台，结合互动教学软件 Nearpod 组织课堂活动进行即时互动。目前诸多的网络语言教学项目中常用的一些在线视频工具，与 Zoom 相似的平台还有 WizIQ、Adobe Connect、Google Hangouts、GoToMeeting 等。与同类的视频工具相比， ZOOM 的优点是安装文件很小，界面直观，操作简单易用，并且服务器稳定，分享桌面也很方便。美中不足则是 ZOOM 平台自带的辅助工具功能有限，无法满足语言课堂中设计多样化的互动学习任务的需求。因此，在 ZOOM 平台的基础上加入 Nearpod 这种类似加强型幻灯片的多媒体应用程序就能很好的补足 ZOOM 的局限，另外 Nearpod 还实时监控互动情况，收集、分享和反馈等功能。例如：在学生用 Nearpod 完成课堂检测时，教师可以通过教师界面的应用程序来实时掌握学生的答题进度。除此之外，教师在利用 Nearpod 设计一些具有针对性的随堂投票或者调查时，可以在收集到的学生的反馈中根据教学需要选定分享相应的内容，进而引入和组织课堂讨论。Nearpod 的这些多样化的功能结合 ZOOM 提供的稳定在线平台大大加强了师生之间、生生之间、以及师生与科技平台之间的课堂互动，使得学生在有限时间的线上同步教学中更好地集中注意力，在密集的互动中最大化目的语的输出和输出质量。

## 4. ZOOM[7]和 Nearpod[8]在课堂互动中的应用

本节将通过具体实例重点讨论 2018 年弗吉尼亚大学星谈项目如何有机结合在线视频工具 ZOOM 和教学互动软件 Nearpod 设计与组织网络中文课堂的教学活动，通过师生、生生与平台之间的多维度互动增强学生在网络语言课堂中的能动性，优化学生的语言输出质量，从而进一步在兼顾趣味性的同时，提高在线语言教学的效率。

### 4.1 ZOOM 和 Nearpod 软件介绍

#### 4.1.1 ZOOM 软件介绍

ZOOM 作为全球大型的在线会议平台软件已经被众多商业和教育机构应用。在 2016 年和 2017 年，弗吉尼亚大学星谈项目先后采用 WizIQ 和 Adobe Connect 作为在线中文教学平台，尽管在功能方面这两个网络教学视频工具与 ZOOM 相比没有太大的不同，然而在实际操作的过程中，ZOOM 在平台稳定性和操作的便捷度方面突显出了优势。因为具有极高的稳定性。因此，在 2018 年，项目决定使用在线视频工具 ZOOM 为网络中文课堂的在线平台。

[7] c.f. https://zoom.us/

[8] c.f. https://nearpod.com/

在用 ZOOM 时，每位老师需要注册一个账号，学生加入老师的网络课堂则不需要注册。如果第一次使用，没有安装 ZOOM 桌面客户端，页面会提示并自动下载 Zoom 客户端。如果用官网 zoom.us 操作, 在网页顶部右侧有三个导航菜单链接：计划（Schedule a Meeting）, 加入（Join a Meeting）, 建立（Host a Meeting）。 图 8 是 ZOOM 官网主页导航菜单。

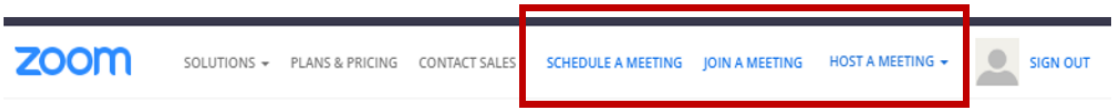



**图 8 ZOOM 官网网页导航菜单**

当安装了 Zoom 客户端之后，客户端主界面(Home) 可以通过计划（Schedule）安排网上课堂的时间，通过“开始”（Start With Video Start Without Video） 在上课开始前进入自己的网络课堂。图 9 是 ZOOM 登录客户端 App，老师可以计划、加入和建立网络课堂，在设置（Settings）中可设置常规选项、音频、视频、录制等。

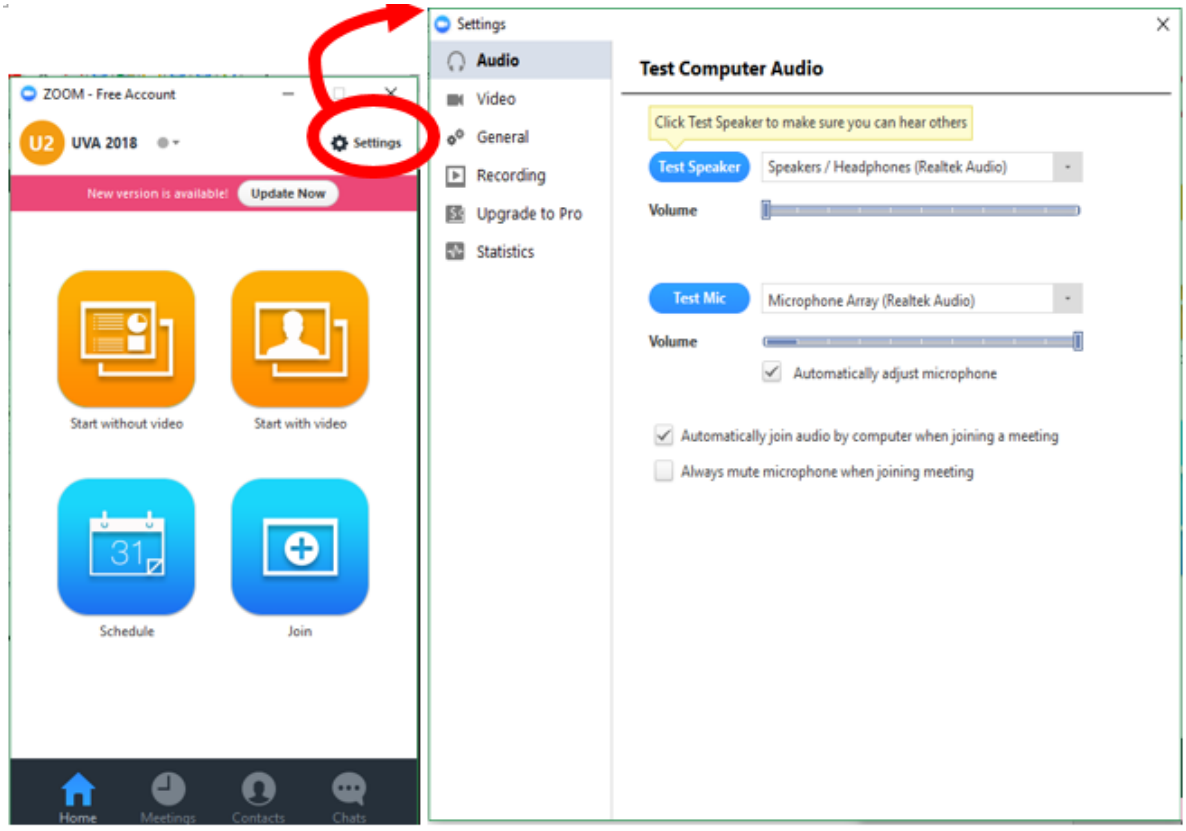



**图 9 ZOOM 登录客户端 App**

进入 ZOOM 网络教室后界面主要有菜单、工具栏、视频或共享主窗口三部分（图 10）。在 ZOOM 网络教室中，教师可以根据实际情况进行多想操作，比如设置音频、视频、邀请其他人、屏幕共享和录制视频等。


**图 10 ZOOM 网络教室操作界面**

学生要想加入 ZOOM 网络课堂两种方式都可以：1. 客户端主界面(Home)，点击 Join，输入 Meeting Id. 学生也可以通过老师发给他们的邀请链接，例如 https://zoom.us/j/58987800，最后的数字即 Meeting Id. 点击链接后，一般系统会用浏览器打开，然后页面提示下载安装或启动已有的客户端程序，点击允许(Allow)继续。这样就无需注册自动加入了。已安装客户端的，可以采用方法点击 Join 然后输入链接最后的数字（Meeting id)。

ZOOM 的也包括一系列辅助功能，例如标注、文字输入、即时聊天和屏幕共享等等，这些辅助功能也能够有效帮助教师在教学活动中与学生进行快速即时、清晰直观增强课堂的互动。在上课的过程中老师也可以点击底部工具栏 Chat，打开聊天窗口，可向所有学生(Everyone)或者指定学生发消息或者教学指令。在线上课程中，最重要的一个功能就是共享屏幕（图 11），老师可以与学生共享电脑主屏幕、白板、平板或手机，也可以单独分享上课的小程序（如打开的幻灯片软件）进行课上的教学活动。


**图 11 ZOOM 共享屏幕的功能**

当开始屏幕分享后，在工具栏中即刻会出现标注（Annotate）选项，老师可以打开 标注（Annotate）选项开始使用标注工具栏的功能（图 12），通过标注工具栏老师与学生可以同时通过打字或者画线在分享的画面上进行互动。以上介绍的是免费账号的功能，如果是付费账号老师还可以通过 breakout 的功能将学生分组进行小组讨论。

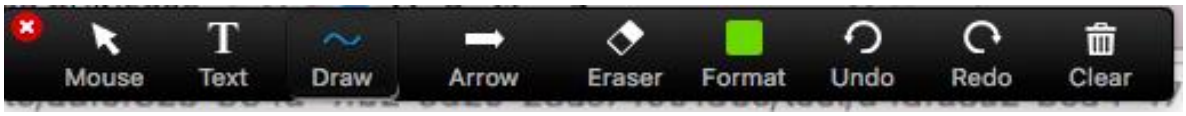



**图 12 ZOOM 标注工具显示图**

有了在线视频工具 ZOOM 提供稳定的教学平台之后，项目的另一个考量是如何针对具体中文教学目标设计课堂语言活动。经过功能多样性、实际操作便捷度、教学素材丰富度等方面的综合考量，此次星谈项目决定在 ZOOM 的基础上结合另一个教学互动软件 Nearpod 进行在线教学的设计。

### 4.1.2 Nearpod 软件介绍

Nearpod 目前已经被广泛地运用在许多 K-12 教学机构中，其丰富多样的活动设计功能在增强课堂教学的趣味性、丰富课堂互动形式上有着不可替代的优势，在语言教学设计上是对在线平台 ZOOM 自带的辅助工具的进一步补充。在本次星谈项目中被师生广泛运用的 Nearpod 活动设计功能包括：在线测验(Quiz)、填空(Fill in the blanks)、连线(Matching)、合作(Collaborate)等等。在网络中文课堂中，教师根据语言点难易程度和教学目标的不同，对 Nearpod 活动进行选择。图 13 为 Nearpod 活动形式显示图。

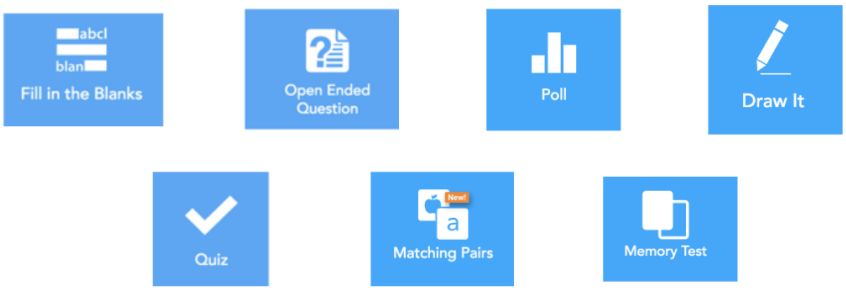



**图 13 Nearpod 活动形式显示图**

在网络课堂进行的过程中，项目推荐学生和教师在电脑界面登陆 ZOOM，进入网络课堂，将电脑屏幕作为课堂的“主屏幕”，同时使用来自平板电脑或智能手机提供的“第二块屏幕”登录教师提前设计好的 Nearpod 活动页面，在两块屏幕共同搭建的平台上参与课堂活动。

## 4.2 课堂互动

在这一节中，文章主要通过师生互动和生生互动两种活动形式介绍 2018 年弗吉尼亚大学星谈项目如何结合 ZOOM 和 Nearpod 设计和组织课堂教学活动。

### 4.2.1 师生互动

师生互动的有效性无论在传统中文课堂亦或网络中文课堂都尤为关键。而以往的线上教学遇到的一大挑战是如何将传统课堂中基于面对面进行的师生互动模式转换到网络平台并保证师生间的教学互动的即时、有效性。在这一过程中，科技辅助工具对于保证实现线上教学的互动质量无疑起着举足轻重的作用。

在本次星谈项目中，以在线视频工具 ZOOM 自带的标注（Annotation）、文字输入等功能为例，教师运用这些简便易操作的辅助功能就可以设计一系列从词句输出到语段输出循序渐进的教学活动。

例如图 14，在学习沉浸式主题“参观北京四中”时，教师可以先罗列出在参观学校设施及讨论学校生活的场景下所需涉及的重点词语/词组，比如：“操场”、“游泳池”、“踢足球”等，要求学生通过 ZOOM 的标注功能将学校的设施与所对应的活动即时连线，并且在教师的引导下一边连线一边用短句来描述“你在……（某个设施/场地）可以做什么？”这是事实层面的语言考察（Facts check）；基于事实层面的问答过后，可以进一步联系到学生个人，要求他们将上述场景中提到的一些活动填在时间关联词后，描述他们在“北京四中”下课后想要做的事，此时学生需要输出的是有实际交际意义(communicative)的句群，而不再是简单的事实层面的单句。ZOOM 的标注功能非常直观地将学生的思路可视化，有效地帮助他们梳理所要表达的内容，并且在教师的引导下一步步扩展语言输出的内容。


**图 14 ZOOM 辅助功能（标注连线功能）**

另一个师生互动的活动则使用到了 ZOOM 辅助功能中的“文字输入”功能。在图 8 中，教师先通过 ZOOM 的 “文字输入” 功能，通过简单的问答要求学生给出与自己在学校的作息日程有关的具体信息，比如 “每天上几节课？”，“第一节课几点开始？”，“有没有午休？”等，多位学生可以同时在线输入自己的答案，既是对事实层面听力理解的核查、也是为之后学生进行具有交际功能的口头报告整理思路。在学生们各自输入相应的答案后，教师进一步要求学生两人一组，根据对方的答案进行提问和回答。由此，在相同的课堂活动中，师生互动也可以过渡到生生互动。图 15 中红色圆圈的部分显示为学生已输入的答案。

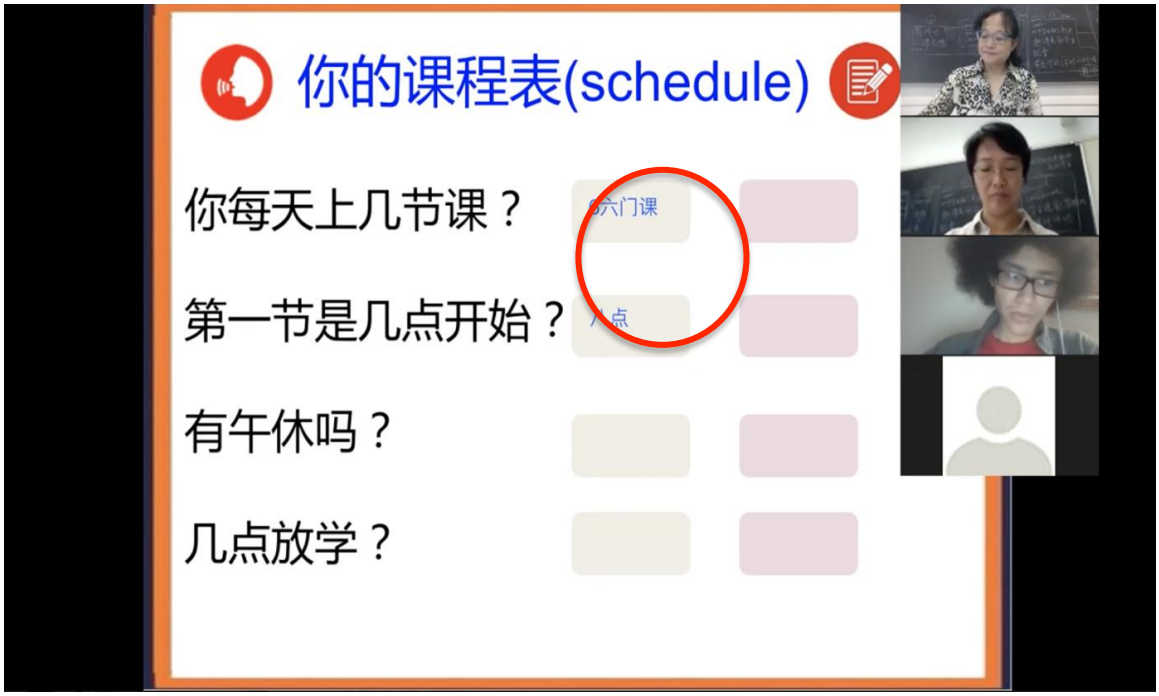



**图 15 ZOOM 辅助功能（文字输入功能）**

由于 ZOOM 视频工具只提供简单的标注和文字输入功能，难以满足教师对课堂活动多样性和延申性的需求，因此，在师生互动部分，项目也引入了 Nearpod 活动。

例如图 16 中展示的是教师通过 Nearpod 中的填空(Fill in the blanks)活动进行的师生互动。教师在课前先根据课文内容和重点语法改写了对话，并针对关键词设计了填空问题。在课堂互动的过程中，学生根据课前语言任务中所学的内容，和老师分别进行角色扮演，并填空完成对话。值得一提的是，这种活动形式也可以在老师和学生的首次配合后转化为生生互动，由两个学生一起完成填空和对话，在学生语言程度较好的情况下，甚至可以进一步开放填空的内容，只给出该交际场景下需要使用的基本句型或关键词，让学生根据词块的提示自行发挥，补充扩展对话内容。

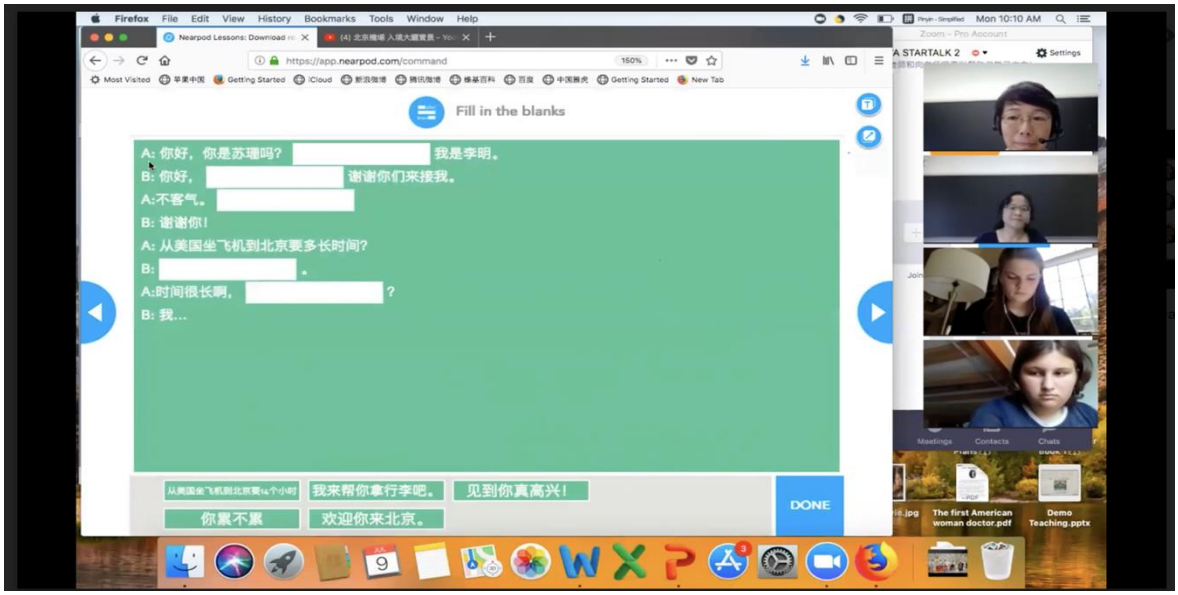



**图 16 Nearpod 活动功能（填空）**

除了填空功能能够辅助课堂师生互动之外，本次星谈项目的教师还使用 Nearpod 的“开放性问题(Open-ended questions)”功能来进行活动。学生通过登陆“第二块屏幕”来回答教师提出的问题，同时在主屏幕上，教师和学生同时能够看到答题情况，也可以随时根据收集到的学生回答进一步互动和补充提问。图 17 中展示的是教师使用开放性问题组织进入课文主题前的词汇复习热身活动的过程。在该活动中，教师提前在 Nearpod 中设置好问题，学生在活动过程中对问题进行回答，每一位提交答案后，教师都可以实时看到答题情况。“开放性问题”的功能主要用来快速、同时收集和共享多位学生的事实性答案，要求学生填写的答案不宜过长，一般以词组、短句为主，因为线上教学的时间宝贵，教师应该有意识地最小化学生的“静默时间（silence time）”，在收集到相关信息后，就应当迅速进入有意义的问答、追问环节，根据学生给出的信息，进一步扩展对话内容。以图 17 中的“饮食习惯”为主题，从“你吃肉吗”的简单是非提问出发，教师可以延展到“你最喜欢/常常吃什么肉？”；“你有没有吃素的朋友？”“吃素/吃肉对身体好不好？”等问题并自然过渡到“饮食/中国菜”的教学主题。如此，学生不仅复习了课前预习的词汇，也始终在进行有意义的语言交流。

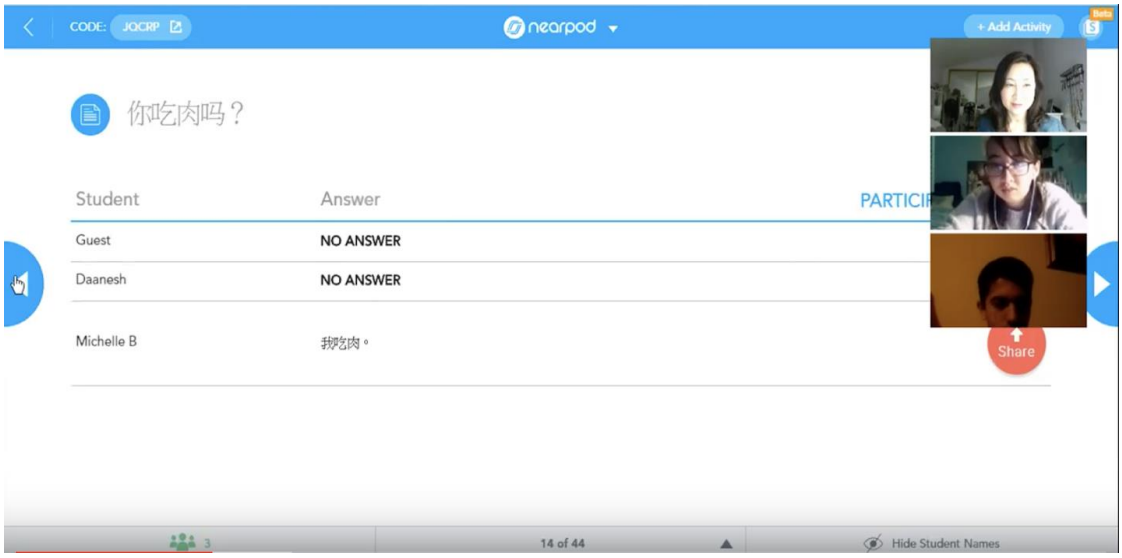



**图 17 Nearpod 活动功能（开放性问题）**

### 4.2.2 生生互动

在星谈项目网络中文课堂上，除了师生互动之外，另一个重要部分就是生生互动。一般认为在传统的课堂中，生生互动的组织和进行显得更为灵活便捷，而在线上教学中，互动模式常常囿于空间的限制而局限在师生的单向互动上。然而，通过本次星谈项目的实践，我们发现借助 ZOOM 视频工具的辅助功能和 Nearpod 多样化的活动形式，教师完全可以将传统课堂中行之有效的生生互动通过合理的设计引导搬上网络课堂。

在 4.2.1 节关于师生互动的部分，我们介绍了教师通过 ZOOM 软件的标注工具“文字标注”介绍学校设施和课外活动的语言任务。在该活动中，教师先在自己和学生之间进行练习，由学生根据要求标注出答案并进行句子及句群层面的问答。师生互动完成之后，则可以使用 ZOOM 辅助功能中的“文字输入”功能，上述课堂活动作进一步延展。通过图 14 中已经完成的第一、第二个语言任务，学生已经对描述学校生活、介绍学校设施所需要的基本词汇、句型比较熟悉，并且能结合自身情况与教师进行有意义的交流对话。接下去的语言任务的目标则是让学生在一个具有真实交际功能的场景下给出成段的输出。于是在图 18 的活动中，教师先给出一个模拟真实生活的场景：作为学生导游带领前来参观学校的国际交换生介绍自己的学校。学生首先需要使用 “文字输入” 的功能在学校的地图上标注出相应的学校场地/设施，之后学生两人一组，一人担当导游的角色，一人扮演交换学生， “导演” 根据自己填入的信息介绍学校的设施和功能，而 “交换学生” 则根据自己听到的信息进行进一步的提问。在这个语言活动中，作为 “导游” 的学生不仅需要给出成段的语言输出，并且与扮演 “交换生” 的学生进行了即时、有实际交际功能的问答。


**图 18 ZOOM 辅助功能（文字输入功能）**

除了 ZOOM 的辅助功能之外，另一个常被项目教师用于生生互动的 Nearpod 活动形式就是“合作活动”(Collaborate)。该活动形式的优点在于学生可以无限次地提交文字信息，针对教师提出的问题和主题分享观点；在提交文字信息之后，教师可以选择性地上传分享,并组织学生进行口头报告。该活动形式因为其较大的自由度而常常作为一节网络中文课堂的收尾活动。在图 19 展示的生生互动中，教师根据“我的家/我家附近”的主题设计了房间陈设对比的活动，在学习完有关房间陈设和与介绍居住环境所需的词汇、语法之后，学生根据自己在美国的家的实际情况对自己居住的地方进行介绍。首先，他们会在手机或平板电脑支持的“第二块屏幕”上输入自己的真实信息，例如家庭住址、房间大小、房型、附近设施等等，然后上传到“合作活动”的活动平台上，教师实时对信息进行选择并发放出来。完成信息的收集、陈列后，学生之间进行相互介绍、报告自己所写的内容，并在教师的引导下对同学给出的信息进行进一步整合处理，比较自己和同学所介绍的居住环境的异同，给出句群到语段的输出，完成生生互动。

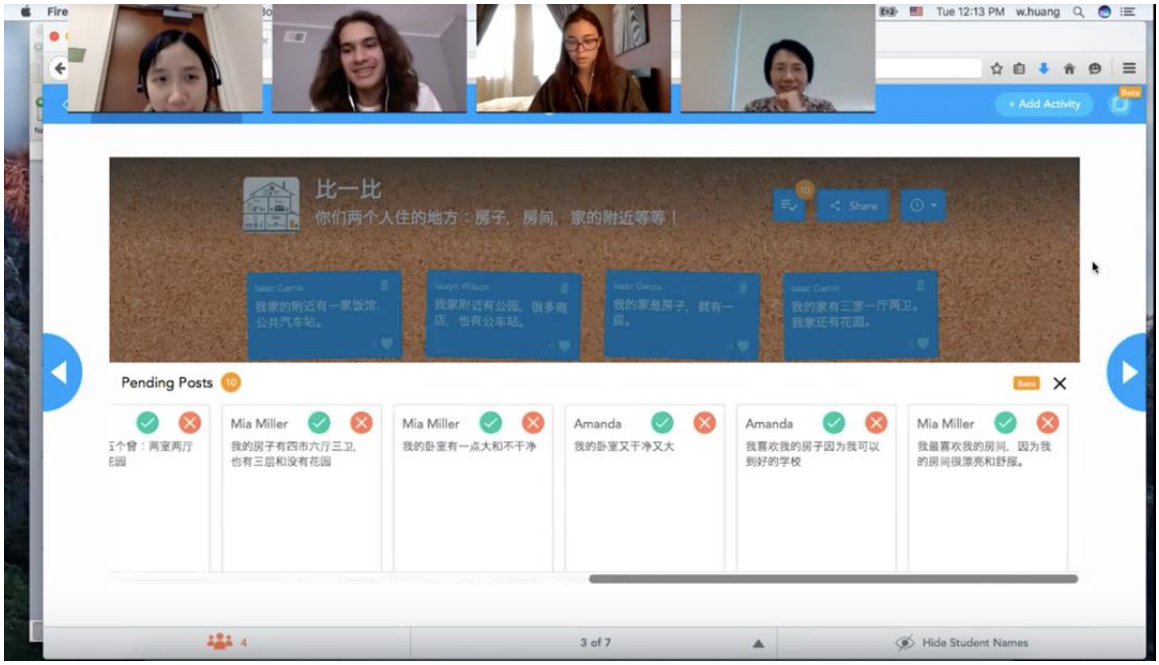



**图 19 Nearpod 活动功能（合作活动）**

结合学生的实际语言水平和学习情况，在星谈项目的网络中文课堂上，教师们通过在线视频工具 ZOOM 和教学互动软件 Nearpod 的活动功能设计了大大小小的互动活动，既弥补了以往线上教学中遭遇的互动性不足的缺陷，又通过层层递进、一个功能多种用途的教学设计方式强化了学生的语言输出质量、同时也增强了课堂的趣味性。

## 5. 结语

弗吉尼亚大学星谈中文项目自 2016 开始致力于运用科技辅助工具将面对面、密集互动的中文教学课堂搬上网络，并在 2018 年开创性地使用了“两块屏幕”—在线视频工具 ZOOM 和互动教学设计软件 Nearpod 相结合的形式来丰富课堂活动、提高语言输出质量。　　在使用新形式和新技术之余，项目也在不懈探索如何在星谈项目的六大“有效教学原则”框架下让科技辅助工具更好地融合进教学设计，更高效地服务于教学目标　　回顾和总结 2018 年星谈项目的成果和经验，在将多种科技辅助工具运动于在线中文教学的实践中，以下几点对教学效果的呈现可说尤为重要：

首先，在以学生为中心的教学设计过程中，科技辅助工具的选取和使用必须时刻考虑学生的体验感，例如：操作是否简便直观、教师给出的指令是否清晰？科技辅助工具可以让教学活动的形式更加多样化，但如何合理安排教学活动的顺序，根据教学目标分配时间，也需要从学生的角度出发做到张弛有度，循序渐进。例如：师生互动性强，简便易操作的小活动可以放在课前热身，而需要学生发挥更多自主性，生生之间需要多步合作的教学任务则可以安排在课程的后半段。另外，密集练习的间隙，引入多媒体真实语料也可以起到让学生耳目一新的作用，同时也为学生营造在真实情境下学习语言及文化的氛围。

其次，在观察和分析教师们实践教学设计的过程中，我们也发现一堂成功的网络中文课未必需要堆砌五花八门的活动，教学辅助工具也不是用得越多、学生的体验和产出越优。在这个过程中，不能忽视的一点是，在目前这一阶段，科技工具还是只是起到辅助语言教学的作用，教师对于教学目标和学生特点的把握仍然决定了他们能否有效地选取和使用对科技工具的选择和使用。具体而言，教师在设计每一个教学活动的时候，都需要考虑如何最大化地利用一个素材 / 一种功能，层层推进学生的产出。例如：以一段多媒体素材作为引入，教师应当设计出有层次的若干个活动，从短句的产出到成段的表达优化语言质量，将学生从事实性问答推向具有交际功能的语言交流。除此之外，还应当注意的一点是，在以听说为主的在线语言教学中，教师需要严格控制课堂中的"静默时间（silence time）"，在选用一些服务于读写的科技辅助工具时，花费的时间不宜过长，且读写的任务需要于听说的任务有机结合，相辅相成。

总而言之，无论是线上教学、混合式教学、或者面对面教学，必须始终以教学目标为最高指导原则，而以科技教学工具为辅，以达成预定教学目标。因此，师资培训千万不能本末倒置，颠倒教学目标与科技工具两者的主辅关系。一言以贯之，师资培训以课程设计与教学有效原则为培训重点，有针对性地进行科技工具评估与其在教学方面的应用与结合，仍然是有效师资培训的总指导方针。最后，必须一提的是，本文着重于介绍 ZOOM 与 Nearpod 如何互相有效结合互动工具以进行线上的中文教学，但其实，ZOOM 与 Nearpod 也适用于混合式教学与面对面两种教学模式，而此一话题非本文重点，有待日后在其它专文中继续讨论。

## 参考文献

Hoopingarner, D. (2009). Best practices in technology and language teaching. *Language and Linguistics Compass*, *3*(1), 222-235.

Warschauer, M., & Meskill, C. (2000). Technology and second language learning. In J. Rosenthal (Ed.), *Handbook of undergraduate second language education* (pp. 303-318). Mahwah, New Jersey: Lawrence Erlbaum.

URL (网址): http://www.tclt.us/journal
Email (电子邮件): editor@tclt.us