

## Chinesisch-Ausspracheunterricht im Distanzmodus: Herausforderungen im Bereich des Perzeptionstrainings

Julia Renner und David Lobner

### Abstract

Der folgende Beitrag setzt sich mit dem Ausspracheunterricht im ersten Studienjahr an der Wiener Sinologie auseinander, der im Wintersemester 2020/21 coronabedingt per Video-Chat stattfinden musste. Der Fokus liegt dabei auf der Umsetzung des Perzeptionstrainings, das im Wesentlichen mittels Pinyin-Diktate erfolgt. Der Beitrag stützt sich einerseits auf eine Fehleranalyse von Pinyin-Diktaten (durchgeführt im synchronen Audio- und Videomodus), andererseits auf eine ergänzende Online-Befragung der Studierenden. Die Ergebnisse zeigen, dass die Perzeption von Anlauten (insbesondere nichtaspirierte) des Mandarin-Chinesischen in der mündlichen, computervermittelten Kommunikation speziell störungsanfällig ist, wobei die Lernenden die Sichtbarkeit von Mund- und Lippenbewegungen, anders als erwartet, größtenteils nicht als deutliche Erleichterung empfinden. Auf Basis der Ergebnisse werden unterschiedliche Pinyin-Diktat-Varianten und ihre jeweiligen Vor- und Nachteile aufgezeigt.

**Keywords:** Distanzunterricht, ChaF-Ausspracheunterricht, Perzeptionstraining, Pinyin-Diktat, sinologische Sprachausbildung

Das sogenannte „Sprachlabor“ im Rahmen des Sinologiestudiums der Universität Wien erstreckt sich über zwei Semester (à 2 Wochenstunden) und fokussiert die gezielte Schulung der Aussprache.

Als Ziel des ersten Semesters wird im Vorlesungsverzeichnis die Beherrschung der „Grundlagen der Aussprache der modernen chinesischen Standardsprache inkl. des Pinyin-Systems“ angeführt. Der Unterricht in Präsenzform gestaltet sich weitgehend durch das Hören und Nachsprechen in der Gruppe mit anschließender individueller Korrektur, wobei jene Personen, die gerade nicht korrigiert werden, an ihrem Computer Aufnahmen hören bzw. selbstständig Aufgaben bearbeiten. Außerdem erfolgt insbesondere im ersten Semester die gezielte Schulung des Gehörs durch Pinyin-Diktate.

Im Wintersemester 2020/21 wurde das Sprachlabor als hybrid-digitaler Unterricht konzipiert. Dies bedeutete konkret, dass der Unterricht zu Beginn des Semesters (Anfang Oktober 2020) als „hybrider“ Unterricht startete und im Verlauf des Semesters, aufgrund der gestiegenen COVID19-Infektionszahlen, als rein digitaler Unterricht weitergeführt wurde.

Die Umsetzung der Hybridform erfolgte in unserem Fall durch die Teilung unserer beiden Lehrveranstaltungsgruppen in jeweils zwei Gruppen. Die Kontaktzeit in Präsenzform betrug pro Gruppe zwar nur die Hälfte, konnte jedoch effizient genutzt werden, da gezielte Vor- und Nachbereitung sowie Kompensationsübungen in asynchroner Form angeleitet wurden. Die digitale Form des Sprachlabors wurde zur regulären Unterrichtszeit über einen synchronen Video-Chat (Blackboard (BB) Collaborate) abgewickelt.

Aufgrund der Forderung, in diesem Semester klassische Prüfungsszenarien zu reduzieren und dafür Formen der laufenden Leistungsbeurteilung mehr Gewicht zuzumessen, wurde folgende Notenzusammensetzung gewählt: 50% der Leistung wurde durch die Erstellung von Audioaufnahmen erbracht, 30% durch die Absolvierung von Pinyin-Diktaten und bis zu 20% nahm die potenzielle Mitarbeit während des synchronen Onlineunterrichts ein.

## **Audioaufnahmen**

Wesentliche Grundlage der Leistungsbeurteilung stellten Audioaufnahmen der Studierenden dar. Bereits im Sommersemester 2020 erprobt ist dieses Verfahren unserer Ansicht nach äußerst effektiv, da die Aufnahme im Vergleich zum gesprochenen Wort wiederholt abspielbar ist und so den Lehrenden noch gezielter ermöglicht, die Aussprache zu analysieren und Rückmeldungen zu geben. Außerdem kann eine weitgehend störungsfreie Akustik gewährleistet werden. Studierende haben wiederum die Möglichkeit, nach dem 1:1-Feedback auf die Aufnahme zurückzukommen und an ihren Fehlern selbstständig weiterzuarbeiten.

## **Pinyin-Diktate**

Insbesondere im ersten Semester der Sprachausbildung an der Wiener Sinologie stellen Pinyin-Diktate eine wesentliche Übungsform dar, um das Gehör für Laute und Töne des Mandarin-Chinesischen zu trainieren.

Wie bereits vermutet<sup>1</sup> stellte die Übertragung von Pinyin-Diktaten in ein digitales Format eine Herausforderung dar, da unterschiedliche Online-Varianten mit spezifischen Vor- und Nachteilen verbunden sind, die abgewogen werden müssen.

In unseren Lehrveranstaltungen wurden zunächst die Formate „Audio (synchron)“ und „Video (synchron)“ erprobt. Diesbezüglich wurde von den Studierenden rückgemeldet, dass die Audioqualität im Audiomodus weniger störungsanfällig sei als im Videomodus des Chats. Als Alternativform zum „live“ Pinyin-Diktat über Audio-/bzw. Video-Chat wurde eine vorab aufgenommene Audiodatei hochgeladen („Audio (asynchron)“), die für alle Stu-

---

<sup>1</sup> Siehe Kurzbericht zum digitalen Sprachlabor (SS2020), der Anfang September 2020 über den Verteiler des Fachverbandes an alle Mitglieder ausgesandt wurde.

dierenden für einen konkret festgelegten Zeitraum freigegeben wurde. Diese Variante erhielt durchwegs positive Rückmeldungen hinsichtlich der Tonqualität, die fehlende Sichtbarkeit der Mundbewegungen wurde allerdings von einigen Studierenden als Manko ins Treffen geführt. Von einer Variante über „Video (asynchron)“ wurde aufgrund der großen Dateigröße und der damit verbundenen Problematik der längeren Downloadzeit vorerst abgesehen.

Um Aufschluss über die Realisierung von Pinyin-Diktaten in unterschiedlichen computervermittelten Modi geben zu können, werden im nachfolgenden Abschnitt die Varianten „Audio (synchron)“ und „Video (synchron)“ mittels einer Fehleranalyse evaluiert. Im Anschluss daran erfolgt die Darstellung einer Kurzbefragung der Studierenden, die thematisch darauf Bezug nimmt.

## Fehleranalyse von Pinyin-Diktaten im digitalen Ausspracheunterricht

Die Fehleranalyse wurde anhand von insgesamt 41 Pinyin-Diktaten, die von zwei Studierendengruppen stammen, durchgeführt. Gruppe 1 bestand aus 16 Studierenden, Gruppe 2 aus 25 Studierenden. Die Beispiele für das Diktat wurden von den zwei LehrveranstaltungsleiterInnen individuell zusammengestellt, wobei vorab die Kriterien festgelegt wurden (15 Silben, jeweils einsilbig; variierende Töne; Fokus auf den Anlauten „j“, „q“, „x“, „zh“, „ch“, „sh“, „z“, „c“, „s“, „r“). Beide Diktate wurden zunächst über BB Collaborate im Audiomodus (ohne Video) durchgeführt:

**Tabelle 1: Pinyin-Diktat Gruppe 1**

chāo xìng jiǎn cuì qióng shī zhè zé xùn chǎn sōng rè zāng xuè shuáng
---

Quelle: Eigene Zusammenstellung.

**Tabelle 2: Pinyin-Diktat Gruppe 2**

zāng rè qióng yán xìng jiā cù qǔ cán xiào chǎo zōng zì cí sān
--

Quelle: Eigene Zusammenstellung.

**Tabelle 3: Fehleranalyse Audio: Gruppe 1**

Silbe	Anlautfehler
chāo	zh (1)
xìng	s (1); zh (1)
jiǎn	d (4); q (1)
cuì	z (5); s (2); ts (1); t (1); n.a. (1)

Silbe	Anlautfehler
qióng	ch (4); j (3); x (2); t (1)
shī	x (2)
zhè	<b>d (2x); g (2)</b>
zé	<b>d (4); c (2); zh (1); r (1)</b>
xùn	sh (4); n.a. (1)
chǎn	zh (2); k (1)
sōng	ch (1); c (1)
rè	j (1); zh (1)
<b>zāng</b>	<b>d (3); b (1); c (1); s (1); zh (1)</b>
xuè	sh (1)
shuáng	ch (1)

Quelle: Eigene Zusammenstellung.

**Tabelle 4: Fehleranalyse Audio: Gruppe 2**

Silbe	Anlautfehler
<b>zāng</b>	<b>d (3); t (3); c (1); zh (1); ch (1); j (1)</b>
rè	sh (1); zh (1)
qióng	ch (8); j (2); x (1)
yán	r (2); i (1)
xǐng:	sh (2); ch (1); q (1)
jiā	q (2); c (2); x (1); z (1); n.a. (1)
cù	z (4); zh (2); dz (1); j (1); n.a. (2)
qǔ	ch (5); j (3); zh (2); q (1)
<b>cán</b>	zh (3); z (2); <b>t (1)</b> ; ch (1)
xiào	y (1)
chǎo	zh (3); j (2); k (1); q (1); c (1); t (1)
<b>zōng</b>	ds (2); zh (2); c (2); <b>d (1)</b> ; t (1); q (1)
zì	c (4); zh (1)
cí	z (3); s (1); x (1)
sān	sh (1)

Quelle: Eigene Zusammenstellung.

Die erste Analyse ergab, dass Anlautfehler die größte Fehlerquelle darstellten. Aus diesem Grund bezieht sich die weiter fokussierte Analyse ausschließlich darauf. Auch Rückmeldungen der Studierenden zufolge (siehe nachfolgende Befragung) wurde die Perzeption von Anlauten als besonders stark durch den computervermittelten Kontext beeinträchtigt empfunden. Im Zuge der Analyse kristallisierten sich vier Hauptkategorien von Anlautfehlern heraus:

a) Fehler, die tendenziell weniger auf der Perzeption beruhen, sondern auf der mangelnden Kenntnis der Silbens Schreibweise bzw. mangelnder Kenntnis des Silbeninventars: z. B. *dsōng* statt *zōng*; *tsuì* statt *cui*; *jè* statt *rè*.

b) „Typische“ Differenzierungsfehler, die tendenziell auf der Perzeption beruhen und aus dem Präsenzünterricht bekannt sind. Diese Kategorie bildet

die größte Gruppe an Anlautfehlern: z. B. Verwechslung von aspirierten und nichtaspirierten Anlauten aus den Anlautgruppen „j, q, x“ – „zh, ch, sh, r“ – „z, c, s“: z. B. *zù* bzw. *sù* statt *cù*; *zù* bzw. *zhù* statt *cù*; *zí* bzw. *sí* statt *cí*.

c) Fehler, die tendenziell auf einer Kombination aus mangelnder Kenntnis der Silbenschriftweise bzw. mangelnder Kenntnis des Silbeninventars und Perzeptionsschwierigkeiten beruhen (Kombination von Kategorie 1 und 2): z. B. *chióng* statt *qióng*.

d) „Untypische“ Differenzierungsfehler (hervorgehoben in der Tabelle), die nicht aus dem Präsenzunterricht bekannt sind und daher tendenziell auf die computervermittelte Modalität zurückzuführen sind. Unter diese Kategorie fallen Anlautfehler, die unserer Unterrichtserfahrung nach im Präsenz-Ausspracheunterricht keine bzw. nur eine untergeordnete Rolle spielen, im Unterricht über Audio- bzw. Video-Chat aber auffallen. Es handelt sich dabei speziell um nichtaspirierte Anlaute, die mit verschiedenen nichtaspirierten Anlauten verwechselt werden: z. B. *diǎn* statt *jiǎn*; *dè* bzw. *gè* statt *zhè*; *dé* statt *gé*; *dāng/bāng/tāng* statt *zāng*; (in beiden Gruppen); *dōng* bzw. *tōng* statt *zōng*; in Einzelfällen traten auch bei aspirierten Silben wie z. B. *cán* oder *cui* untypische Anlautfehler auf (Verwechslung mit „t“).

Die Ergebnisse des Pinyin-Diktats wurden im Anschluss mit den Studierenden gemeinsam besprochen. Als didaktische Maßnahme wurde in Gruppe 1 in der darauffolgenden Stunde dasselbe Diktat nochmals im Videomodus durchgeführt, um einen Vergleich mit der Variante mittels Audiomodus zu ziehen und den Studierenden in dieser neuen, herausfordernden Situation eine Verbesserungsmöglichkeit einzuräumen.

Im Vergleich zur Audiovariante konnte in Gruppe 1 die Fehlerquote von 26% (63 Fehler von maximal 240 Fehlern / 16 Studis à 15 Silbenbeispiele) auf 18% (44 Fehler von maximal 240 Fehlern / 16 Studis à 15 Silbenbeispiele) reduziert werden. Inwiefern hier die Reduktion der Fehler auf den Faktor „Video“ zurückzuführen ist, kann an dieser Stelle nicht eindeutig bestätigt werden, da es sich bei der Durchführung über Audiochat um einen für die Studierenden unbekanntem Test gehandelt hat und in der Variante über Videochat um die Wiederholung des Tests aus der Vorwoche. Zudem stellte die Testsituation kein kontrolliertes Studiensetting dar. Es besteht daher die Möglichkeit von variierenden technischen, räumlichen und akustischen Rahmenbedingungen.

Die Analyse weist allerdings darauf hin, dass bei verwechselten Anlauten, die eine deutlich sichtbare, unterschiedliche Mund-/Lippen-/Zungenposition aufweisen, die visuelle Komponente eine hilfreiche Rolle bei der Identifikation gespielt hat (z. B. „*zhè*“ – gehört als „*gè*“ in der ersten Runde – wurde in der zweiten Runde korrekt identifiziert; der Anlaut „*c*“, wie in „*cui*“ wurde in der zweiten Runde zumindest nicht mehr mit „*t*“ verwechselt. Hingegen scheint die visuelle Komponente keine hilfreiche Rolle bei der Differenzierung von Anlauten einer Anlautgruppe (wie z. B. z, c, s) zu neh-

men, da hier die Mund-/Lippen-/Zungenposition sehr ähnlich ist. Das Problem, dass die nichtaspirierten Anlaute aus der Kategorie (z, c, s bzw. zh, ch, sh, r) mit dem Konsonanten „d“ verwechselt wurden, blieb im Wesentlichen bestehen und scheint, nach eingängiger Diskussion mit Studierenden und KollegInnen, die Mandarin-Onlineunterricht in Anspruch nehmen bzw. geben, ein spezielles Phänomen zu sein, das mit den akustischen Rahmenbedingungen der computervermittelten Kommunikation zusammenhängt.

**Tabelle 5: Fehleranalyse Video: Gruppe 1**

Silbe	Anlautfehler
chāo	zh (1)
xìng	q (1)
jiǎn	d (1); q (1)
cùi	z (3); s (3)
qióng	j (3); ch (3); x (1)
shī	x (1); s (1); n.a. (1)
zhè	d (2x); z (1); q (1); n.a. (1)
zé	d (3); c (1); ch (1); n.a. (1)
xùn	sh (2)
chǎn	zh (1)
sōng	c (2)
rè	j (2)
zāng	c (3); zh (1)
xuè	sh (1); n.a. (1)
shuáng	0

Quelle: Eigene Zusammenstellung.

## Befragung der Studierenden

Um die Studierenden zu ihren Ansichten zu befragen und in den Entscheidungsprozess bezüglich der weiteren Durchführung der Pinyin-Diktate einzubinden, wurden beide Kursgruppen schriftlich befragt. Die Befragung wurde in einer Gruppe direkt über die „Polling“-Funktion von Blackboard Collaborate durchgeführt, in der anderen Gruppe wurde sie über die Befragungsfunktion von Moodle abgewickelt. Insgesamt nahmen 35 Studierende an der Befragung teil.

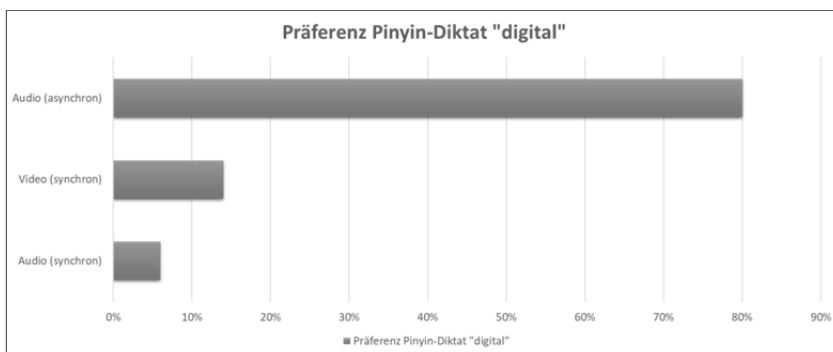
Die Befragung beinhaltete drei Fragen (siehe Anhang) und war als Multiple-Choice-Fragebogen angelegt. Frage 1 bezog sich auf die präferierte digitale Pinyin-Diktat-Variante. Hierbei wurde als Auswahlmöglichkeit neben den bereits erprobten Varianten „Video (synchron)“ und „Audio (syn-

chron)“ auf Anregung der Studierenden auch die Variante „Audio (asynchron)“ miteinbezogen.<sup>2</sup>

Frage 2 bezog sich auf die Bedeutung der Sichtbarkeit der Mundbewegungen für die Perception und Frage 3 auf empfundenen Beeinträchtigungen aufgrund des computervermittelten Kommunikationskontexts.

Hinsichtlich der bevorzugten digitalen Pinyin-Diktat-Variante zeigten sich folgende Präferenzen: Von insgesamt 37 ausgewählten Antworten entfielen 80% auf jene Variante mit der vorab aufgenommenen Audio-Datei, 14% auf die Video-Chat-Variante und 6% auf die Audio-Chat-Variante.

**Abb. 1: Digitale Pinyin-Diktat-Variante**

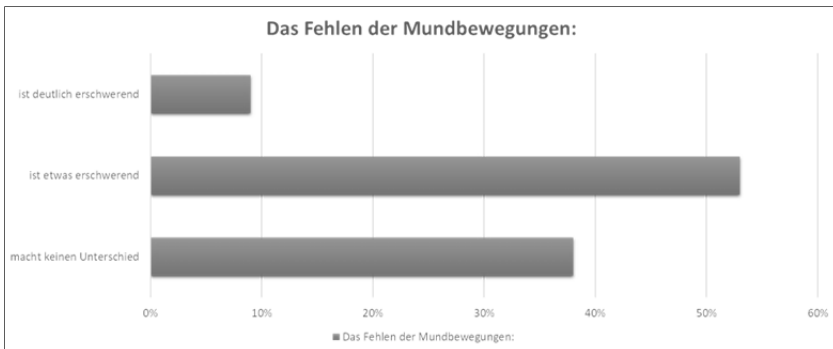


Quelle: Eigene Berechnung und Zusammenstellung.

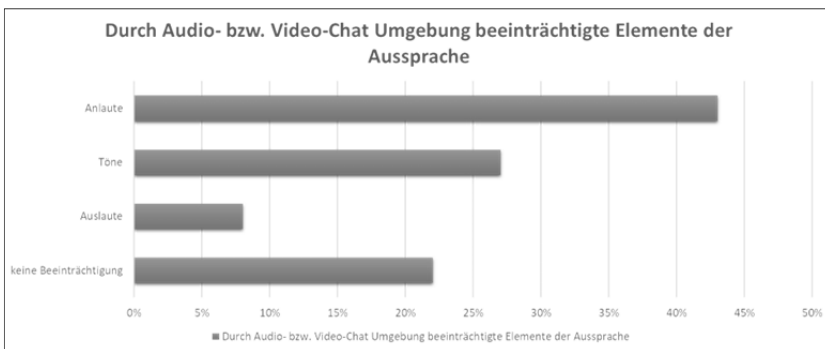
Hinsichtlich der Bedeutung der Sichtbarkeit der Mundbewegungen zeigt sich, dass von 34 ausgewählten Antworten 9% das Fehlen als „deutlich erschwerend“ und 53% als „etwas erschwerend“ einstufen. 38% sahen im Wegfall der visuellen Komponente keinen Nachteil.

Des Weiteren zeigte sich, dass insbesondere Anlaute durch den Audio- bzw. Video-Chat-Modus als undeutlich empfunden werden. Von insgesamt 40 ausgewählten Antworten entfielen 42,5% auf „Anlaute“, 27,5% – entgegen unserer Erwartungen – auf „Töne“, 7,5% auf „Auslaute“. 22,5% der Antworten entfielen auf die Kategorie „keine Beeinträchtigung“.

<sup>2</sup> Die Variante „Video (asynchron)“ wurde aufgrund der großen Dateigröße, die eine derartige Aufnahme mit sich bringen würde, ausgeschlossen. Da den Studierenden für die asynchrone Variante nur ein Zeitfenster von 10 Minuten (inkl. Hochladephase) gewährt wurde, schien der Download eines Videos als zu störungsanfällig.

**Abb. 2: Sichtbarkeit der Mundbewegungen**

Quelle: Eigene Berechnung und Zusammenstellung.

**Abb. 3: Beeinträchtigungen im Audio- bzw. Video-Chat-Modus**

Quelle: Eigene Berechnung und Zusammenstellung.

## Pädagogische Schlussfolgerungen

Da die überwiegende Mehrheit für eine Durchführung im asynchronen Modus mittels Audioaufnahme plädierte, wurde von beiden LehrveranstaltungsleiterInnen für das verbleibende Semester diese Variante übernommen. Rückmeldungen der Studierenden zufolge empfanden weniger als 10% aller Studierenden den Wegfall der visuellen Komponente bei Pinyin-Diktaten als „deutlich erschwerend“. Diskussionen im Anschluss an die Befragung ergaben, dass die Stabilität der Akustik, die eine vorab aufgenommene Audiodatei mit sich bringt, den Studierenden die notwendige Sicherheit gibt, die als wichtiger als die visuelle Komponente gesehen wird.



Auf Basis unserer Untersuchung und unserer Lehrerfahrung empfehlen wir für die Durchführung von Pinyin-Diktaten im Onlineunterricht folgende Vorgehensweise:

- Hochladen der Audioaufnahme auf Lernplattform (Zusendung per Mail wurde vermieden, damit eine zeitversetzte Übermittlung vermieden werden kann) mit gleichzeitiger Bekanntgabe eines konkreten Bearbeitungszeitfensters.
- Optional: Anschließendes Durchgehen der einzelnen Items im Video-Chat, um die visuelle Komponente zu „ergänzen“.
- Finale Abgabe.

Inhaltlich wurde, anders als bei Pinyin-Diktaten im Präsenzunterricht, wo in der Regel unbekanntes Vokabular als Basis dient, auf eine Mischung von bekanntem Vokabular aus den Unterrichtslektionen und unbekanntem Vokabular zurückgegriffen. Die Anzahl der Silben wurde zudem sukzessive von einsilbigen Beispielen bis zu mehrsilbigen Chengyus erweitert. Auch das Sprechtempo wurde mit der Zeit erhöht.

Nachfolgende Tabelle fasst in kompakter Form nochmals die jeweiligen Vor- und Nachteile der einzelnen Modi für Pinyin-Diktate zusammen, um KollegInnen einen schnellen Überblick zu verschaffen:

**Tabelle 6: Überblick der unterschiedlichen Pinyin-Diktat-Modi**

<b>Modus</b>	<b>Vorteile</b>	<b>Nachteile</b>
Audio (synchron)	Tonqualität besser bzw. weniger störungsanfällig als über Video-Chat störungsanfälliger als über Audio- und Videoaufnahmen Beginn und Ende durch die Lehrperson „live“ kontrollierbar Möglichkeit „live“ Nachfragen zu stellen	keine Mundbewegungen sichtbar Tonqualität schlechter bzw. störungsanfälliger als asynchrones Audio und Video
Video (synchron)	Mundbewegungen sichtbar Beginn und Ende durch die Lehrperson „live“ kontrollierbar Möglichkeit „live“ Nachfragen zu stellen	Tonqualität störungsanfälliger als bei allen anderen Formen (audio-synchron; audio-/video-asynchron)

<b>Modus</b>	<b>Vorteile</b>	<b>Nachteile</b>
Audio (asynchron)	Sehr gute Tonqualität Nicht störungsanfällig Geringe Downloadzeit, da kleine Dateigröße	Keine Mundbewegungen sichtbar Beginn und Ende durch die Lehrperson nicht 100% kontrollierbar
Video (asynchron)	Sehr gute Tonqualität Nicht störungsanfällig Mundbewegungen sichtbar	Beginn und Ende durch die Lehrperson nicht 100% kontrollierbar Längere Downloadzeit, da größere Dateigröße

Quelle: Eigene Zusammenstellung.

## **Ausblick**

Die empirische Studie des Beitrags entstand im Zuge unserer Überlegungen und Beobachtungen bei der Umsetzung von Pinyin-Diktaten im digitalen Chinesischunterricht und wurde nicht in erster Linie als Forschungsprojekt realisiert. Das Studiendesign und der Umfang der Ergebnisse sind daher vor diesem Hintergrund zu betrachten. Nichtsdestotrotz gaben die Ergebnisse Hinweise darauf, dass bestimmte Anlaute des Mandarin-Chinesischen (insbesondere nichtaspirierte) in der mündlichen, computervermittelten Kommunikation offensichtlich speziellen Schwierigkeiten unterliegen, die mit den akustischen Rahmenbedingungen zusammenhängen. Zukünftige Studien könnten daher in einer kontrollierten Herangehensweise die Perzeption im online-basierten Chinesischunterricht untersuchen, um Aufschluss über potenziell systematische Störfaktoren geben zu können.

## **Anhang Online-Fragebogen**

### **In der digitalen Variante des Pinyin-Diktates präferiere ich:**

Audio (synchron über Collaborate)

Video (synchron über Collaborate)

eine vorab aufgenommene Audiodatei

### **Die Mundbewegungen der Lehrkraft beim Pinyin-Diktat nicht zu sehen empfinde ich als**

deutlich erschwerend

etwas erschwerend

macht für mich keinen Unterschied

**Im Unterricht über Collaborate sind für mich folgende Aspekte der Aussprache besonders beeinträchtigt:**

Anlaute

Auslaute

Töne

keine Beeinträchtigung

**Abstract**

This article focuses on perception training through “pinyin dictations” in an online pronunciation course for learners of Mandarin as a foreign language. Three different modes (synchronous audio; synchronous video; asynchronous audio) are compared through an error analysis and an online survey of the students. The results show that the perception of syllable onsets, especially non-aspirated sounds, is particularly disturbed through the synchronous computer-mediated environment. Contrary to expectations, the visual component (visibility of movement of the mouth and lips) was not perceived as crucial. Based on the results the advantages and drawbacks of each pinyin dictation variant are discussed.

Manuskript eingereicht am 28.02.2021; akzeptiert am 30.04.2021