

Jürgen Trouvain\*

# Das IFCASL-Korpus als phonetisches Lernerkorpus

## The IFCASL corpus as a phonetic learner corpus

<https://doi.org/10.1515/zgl-2022-2049>

**Abstract:** Phonetic learner corpora represent a special type of spoken learner corpora by providing detailed phonetic and phonological annotation in the form of phonetic transcription as well as the segmentation and labelling of the speech signal on the levels of segments, syllables, words and sentences. This time-consuming post-processing enables a better acoustic analysis of the data and provides many options for using audio data in teaching foreign languages. It also offers good opportunities to develop tools for computer-aided pronunciation training. The possibilities and restrictions of phonetic learner corpora are illustrated with the bidirectional German–French IFCASL corpus.

- 1 Einführung
- 2 Beschreibung des IFCASL-Korpus
  - 2.1 Teilnehmer
  - 2.2 Sprechaufgaben
  - 2.3 Subkorpora
  - 2.4 Annotationen
- 3 Anwendungsbeispiele
  - 3.1 Messphonetische Studien der Produktion und Perzeption
  - 3.2 Lehre von DaF und in der Ausbildung von DaF-Lehrkräften
  - 3.3 Entwicklung von computergestütztem Aussprachetraining
- 4 Beschränkungen
  - 4.1 Fehlende Spontansprache
  - 4.2 Instrumentelle Erweiterungen
  - 4.3 Datenmanagement
  - 4.4 Einsatz für Lehrkräfte
- 5 Zusammenfassung und Ausblick  
Danksagung  
Literatur

---

**\*Kontaktperson:** Dr. Jürgen Trouvain: Universität des Saarlandes, Fachrichtung Sprachwissenschaft und Sprachtechnologie, Campus C7.2, D-66123 Saarbrücken, E-Mail: [trouvain@lst.uni-saarland.de](mailto:trouvain@lst.uni-saarland.de)

# 1 Einführung

Ein Überblick zu Lernerkorpora wie bei CECL (2021) zeigt zwei Auffälligkeiten: 1) die meisten Lernerkorpora bestehen aus Daten mit geschriebener Sprache und 2) weisen Englisch als Zielsprache auf. Auch wenn beides nicht überrascht, so lohnt es sich bewusst zu machen, dass Lernerkorpora eher selten gesprochene Sprache beinhalten (Ballier & Martin 2015, Trouvain et al. 2017), obwohl mündliche Kommunikation für Lerner (= L2-Sprecher) genauso wie für Muttersprachler<sup>1</sup> (= L1-Sprecher) eine zentrale Rolle beim Austausch sprachlicher Information darstellt und viele Kommunikationssituationen durch Mündlichkeit und weniger durch Schriftlichkeit geprägt sind.

Es lassen sich unterschiedliche Kategorien gesprochener Lernerkorpora (gLK) unterscheiden, wie unten genauer ausgeführt wird. Die aus einer phonetischen und phonologischen Warte aufbereiteten gLK, die in diesem Aufsatz kurzerhand phonetische Lernerkorpora (phonLK) genannt werden, gehören dabei zu solchen gLK, die seltener zu finden sind. Der vorliegende Aufsatz versucht daher die Besonderheit phonLK zu beleuchten. Möglichkeiten und Begrenzungen von phonLK werden am Beispiel des IFCASL-Korpus<sup>2</sup> (Trouvain et al. 2016) illustriert.

Gesprochene Lernerkorpora (gLK) können aus einer phonetischen Perspektive verschiedenartig unterteilt werden. In ihrem Überblicksaufsatz über die Annotation gesprochener Sprache in Lernerkorpora nehmen Ballier & Martin (2015) eine grobe Einteilung in drei Kategorien vor: 1. „stumme“ („mute“) Lernerkorpora, 2. „gesprochene“ („spoken“) Lernerkorpora und 3. „phonetische“ („phonetic“) Lernerkorpora. In „stummen“ *Lernerkorpora* liegen die Daten ausschließlich als Sekundärdaten in Form von Transkripten vor. Es handelt sich üblicherweise um Spontansprache, die orthographisch verschriftet ist, wobei Marker von Disfluenzen wie Häsitationen und Pausen etc. hinzukommen. Jedoch liegen die zu Grunde liegenden Audiodateien als Primärdaten nicht mehr vor und bilden daher keinen Teil des Korpus. In der Gesprächsanalyse und der interaktionalen Linguistik begegnet man häufig dieser Art von Daten. Ein Beispiel für ein stummes Lernerkorpus ist das LINDSEI-Korpus für L2-Englisch (Gilquin et al. 2010). Im Gegensatz dazu liegen bei „gesprochenen“ *Lernerkorpora* sowohl die

---

1 Zur nicht unproblematischen Abgrenzung zwischen L1 und L2 sei auf die Einführung der Gastherausgeberin und auf Fandrych & Waller (beide in diesem Heft) verwiesen.

2 Mehr Informationen zum IFCASL-Korpus findet man auf <<http://www.ifcasl.org/>>. Das Korpus wird im Laufe des Jahres 2022 für die wissenschaftliche Öffentlichkeit zugänglich gemacht werden.

Transkripte als auch die Audiodateien vor wie beim Berlin Map Task Corpus für L2- und L1-Deutsch (Belz et al. 2017). Idealerweise sind beide Datei-Typen miteinander auf der Satzebene verknüpft. Zumeist handelt es sich um Spontansprache von Lernern (inklusive aufgabenbasierter nicht verschrifteter Sprache), bei denen die Audiodateien als Primärdaten vorliegen und als Sekundärdaten dazugehörige orthographienahe Transkriptionen hinzugefügt sind, mit etwaigen Zusatzannotationen wie Häsitationen, Pausen etc. (siehe dazu Hirschmann & Schmidt, in diesem Heft). „Phonetische“ *Lernerkorpora*, auch zuweilen phonologische Lernerkorpora genannt (Gut 2014), bestehen zunächst aus den Audiodateien als den Primärdaten. Als Sekundärdaten liegen Transkriptionen der aufgenommenen Spontan- und/oder Lesesprache als orthographienahe bzw. orthographische Transkription und in Lautschrift vor. Die Transkripte sind auf unterschiedlichen linguistischen Ebenen zeitlich mit den korrespondierenden Audiodateien aligniert. Annotationen kann es auf Satz-, Wort-, Silben- und Segmentebene geben. Dies erlaubt eine Suche nach Instanzen eines bestimmten Wortes, eines bestimmten Vokals oder Konsonanten bei gleichzeitiger Verknüpfung zum tatsächlichen akustisch aufgenommenen Ereignis der Lernaltersprache. Die zeitalignierte Segmentierung und Annotation erlaubt es, mit einem Speech Editor wie PRAAT (Boersma & Weenink 2016) die jeweiligen zeitlichen Abschnitte des Sprachsignals zu sehen und wiederzugeben, z. B. ein auditives und visuelles Wort-für-Wort-Durchgehen der Audiodatei. Diese Art von Annotation ermöglicht automatische signalbasierte phonetische Analysen und hilft bei manuell durchgeführten phonetischen Sichtungen und Auswertungen. Ein Beispiel für ein solches Korpus ist WroDiaCo (Belz & Odebrecht, in diesem Heft) mit einer automatischen Alignierung auf Wort- und Segmentebene.

Eine alternative Einteilung der gLK ist die Kategorie gLK als Oberbegriff zu betrachten und phonLK als Spezialfall von gLK, so wie es auch Hirschmann & Schmidt (in diesem Heft) vornehmen. Eine Unterscheidung zwischen „stummen“ und „gesprochenen“ LK entfällt somit und damit auch die ungenaue Bezeichnung „spoken“, die zwar einen Gegensatz zu „mute“ ausdrückt, aber nicht zu „phonetic“. In aller Regel enthalten gLK spontan- und nicht lese-sprachliches Material, das transkribiert und ggf. annotiert werden muss. Es sind oftmals die durch Transkription und Annotation gewonnenen Sekundärdaten, die zu weiteren Analysen benutzt werden (siehe alle anderen Beiträge in diesem Heft). In diesen Fällen ist eine Unterscheidung zwischen „mute“ und „spoken“ hinfällig. Der große Vorteil von „spoken“ Korpora ist, dass sie auch die Modalität des Hörens miteinbeziehen und man sehr gezielt ausgewählte Items angehört werden können.

Der Spezialfall phonLK unterscheidet sich oftmals von anderen gLK durch folgende Merkmale:

- die Audioaufnahmen werden in einer optimalen akustischen Umgebung aufgenommen bezüglich 1) der Schall- und Reflexionseigenschaften (z. B. schallbehandelte Aufnahmekabine), 2) der Aufnahmetechnik (z. B. hochwertige Nahsprechmikrophone) und 3) der Vermeidung der akustischen Maskierung durch Überlappen zweier (oder mehrerer) Sprecher (z. B. durch Einzelaufnahmen oder Aufnahmen in (akustisch) getrennten Räumen),
- die Audioaufnahmen werden zu einer phonetischen (und nicht nur orthographienahen) Transkription benutzt,
- die Audioaufnahmen werden einer genauen zeitalignierten Segmentierung und Etikettierung auf Lautebene bzw. auf Silben- und Wortebene unterzogen,
- die Audioaufnahmen werden zu akustisch-phonetischen Analysen benutzt (z. B. zur Messung von Formanten),
- die Aufnahmen können sowohl aus Spontan-, aber auch aus Lesesprache bestehen,
- die Aufnahmen werden gezielt zur Erstellung des Korpus durchgeführt und sind entsprechend konzipiert,
- zu den Primärdaten können neben Audio- und (seltener) Video-Aufnahmen auch physiologische Daten (z. B. Ultraschall) erhoben werden.

Diese Unterschiedlichkeit zwischen phonLK und anderen gLK bedeutet aber nicht, dass nicht auch mit gLK, die nicht als phonLK erstellt worden sind, phonetische und vor allem prosodische Fragestellungen bearbeitet werden können. So kann bei ausreichend guter akustischer Qualität auch in nicht-phonLK Aussprachebesonderheiten von Lernern transkribiert werden. Ebenso können Pausen akustisch (grob) erfasst werden, Disfluenzen können annotiert werden, beides wichtige Aspekte, um Sprechfluss<sup>3</sup> genauer zu untersuchen.

PhonLK sind also ganz gezielt für bestimmte Arten von Daten ausgelegt, die es bei anderen gLK nicht gibt. Der Preis für den Datenreichtum in der phonetischen Tiefe ist aber, dass die Erstellung und Bearbeitung bei phonLK um ein Vielfaches arbeitsaufwändiger und personalintensiver ist als bei anderen gLK: Sämtliche Äußerungen werden auch in Lautschrift transkribiert, ebenso erfordert

---

<sup>3</sup> Sprechfluss bzw. Redefluss wird hier synonym zu Sprechflüssigkeit gebraucht. Bei Fluss wird der Prozess des Sprechens in den Vordergrund gestellt, bei Flüssigkeit das Produkt des Sprechens. Unter Musikern scheint es wohl Usus zu sein, die flüssige musikalische Darbietung als Spielfluss zu bezeichnen, wohingegen Spielflüssigkeit der gesammelte Speichel der Blechbläser bezeichnet, der sich während des Spielens im Instrument ansammelt und zu geeigneten Zeitpunkten entleert wird.

das Alignieren von Wörtern, Silben und Lauten je nach Detailtiefe viel Sorgfalt (Trouvain et al. 2017). Draxler (2008) schätzt den dabei anfallenden Annotationsaufwand auf mehr als 500 Minuten Nachbearbeitungszeit für jede Minute Aufnahme.

Der Rest des Aufsatzes ist wie folgt aufgebaut: In der nachfolgenden Sektion wird ein konkretes phonLK im Detail vorgestellt. Sektion 3 befasst sich mit Beispielen der Anwendung und der Frage wozu die phonetische Genauigkeit notwendig ist. Nach den Möglichkeiten für Forschung und Lehre, die sich durch phonLK bieten, sollen in der vierten Sektion die Beschränkungen vorgestellt werden. Am Schluss erfolgt eine Zusammenfassung und ein Ausblick.

## 2 Beschreibung des IFCASL-Korpus

Die Anwendungsbeispiele in den folgenden Sektionen entstammen alle dem IFCASL-Korpus (Trouvain et al. 2016), das hier in der gebotenen Kürze vorgestellt wird.

### 2.1 Teilnehmer

Für das Korpus wurden 100 Sprecher aufgenommen: 50 mit Deutsch als L1, 50 mit Französisch als L1. Jede der beiden Gruppen besteht aus den folgenden drei Untergruppen (vgl. auch Tab. 2):

20 erwachsene Lerner auf Anfänger-Niveau (nach Selbstauskunft und auf der Skala des Gemeinsamen Europäischen Referenzrahmens für Sprachen GERS: A2, evtl. B1),

20 erwachsene fortgeschrittene Lerner (GERS: B2 und C1),

10 Schüler im Alter von ca. 14 Jahren (GERS: A2).

**Tab. 1:** Anzahl der Sprecher des IFCASL-Korpus für jede der sechs Untergruppen.

	L1	L1
Alter/Niveau	Französisch	Deutsch
Schüler (ca. 14 Jahre)/Anfänger	10	10
Erwachsene/Anfänger	20	20
Erwachsene/Fortgeschrittene	20	20

Die Personenauswahl ist geschlechtermäßig nahezu ausbalanciert. Ein möglicher regionaler Einfluss wurde dadurch ausgeglichen, dass die Sprecher nicht alle aus einer Region Frankreichs bzw. Deutschlands stammen. Die vor den Aufnahmen erhobenen Fragebögen haben ergeben, dass alle Erwachsenen Abitur haben und Englisch als zweite bzw. bevorzugte Fremdsprache verwenden.

## 2.2 Sprechaufgaben

Das gesamte Aufnahmematerial des Korpus ist gelesen. Das vorzulesende Material bestand in jeder der beiden Sprachen aus vier Teilen:

1. *Vorleseaufgabe* mit 31 Sätzen, die orthographisch vorlagen, z. B. *In Berlin zahlt man wenig Miete.*,
2. *Nachsprecheraufgabe* mit 29 Sätzen, die orthographisch und als Audio-Datei vorlagen, letztere wurde von einem L1-Modell-Sprecher vorher aufgenommen (dieser Teil wurde in der L2-Version gewählt, in der L1-Version wurde wie zuvor vorgelesen); für die L2-Sprecher also eigentlich eine lesegestützte Nachsprecheraufgabe, z. B. *Die Oase liegt in der Wüste.*,
3. *Fokusaufgabe* mit sieben Sätzen, deren Fokus sich jeweils ändert, z. B. Frage: *Peter vertraut dem Kamel?* Laut zu lesende Antwort: *Yvonne vertraut dem Kamel.*,
4. *Vorlesegeschichte: Die drei kleinen Schweinchen* (bestehend aus jeweils 13 Sätzen in beiden Sprachen).

Die Durchführung der Aufnahmen dauerte pro Sprecher ca. eine Stunde.

## 2.3 Subkorpora

Das IFCASL-Korpus ist ein bidirektionales Korpus, bei dem alle Sprecher das Vorlesematerial sowohl in ihrer L1 als auch in ihrer L2 produziert haben. Letztendlich liegen also vier Subkorpora vor (vgl. Tab. 2): ein französischsprachiges L1-Korpus (FF), ein deutschsprachiges L1-Korpus (GG), ein französischsprachiges L2-Korpus von Sprechern mit Deutsch als L1 (GF) sowie ein deutschsprachiges L2-Korpus von Sprechern mit Französisch als L1 (FG). Die folgenden Ausführungen beziehen sich in erster Linie auf das letzte Sub-Korpus, wobei auch das deutschsprachige L1-Korpus zuweilen berücksichtigt wird. Ein phonLK bzw. generell ein gLK muss also keineswegs nur Produktionen der Lerner enthalten, sondern kann auch L1-Produktionen der Zielsprache, aber auch L1-Produktionen der Ausgangssprache als Bestandteil haben.

**Tab. 2:** Die vier Sub-Korpora des symmetrisch angelegten IFCASL-Korpus. Von besonderem Interesse für DaF ist das Sub-Korpus mit Sprechern mit L1 Französisch und Deutsch als L2 (FG).

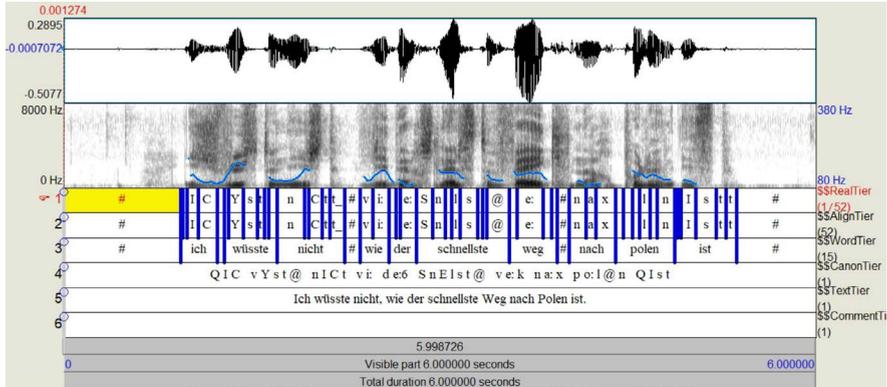
Lesematerial	L1: Französisch	L1: Deutsch
Französisch	FF	GF
Deutsch	FG	GG

## 2.4 Annotation

Für alle Sprachaufnahmen liegen Annotationsdateien in Form von Praat-Text-Grids vor, die beispielsweise mit dem Sprachverarbeitungsprogramm Praat (Boersma & Weenink 2016) angeschaut und weiterverarbeitet werden können. Für die folgenden Ebenen liegen Annotationen vor:

1. Satz in orthographischer Schrift (in Abb. 1 auf Ebene („Tier“) 5),
2. Satz in maschinenlesbarer Lautschrift, z. B. [ʔ η ʃ ç ɪ ʏ ʊ ɔ ε ø œ ɐ ə] = [Q N S C I Y U O E 2 9 6 @] (in Abb. 1 auf Tier 4: [QIC vYst@ nICt vi: de:6 SnElst@ ve:k na:x po:l@n QIst]),
3. alle Wörter (in orthographischer Schrift) mit Wortgrenzen, automatisch mit intern entwickelter Software aligniert (in Abb. 1 auf Tier 3),
4. alle Laute (in maschinenlesbarer Lautschrift) mit Lautgrenzen, automatisch aligniert (in Abb. 1 auf Tier 2),
5. alle Laute (in maschinenlesbarer Lautschrift) mit Lautgrenzen, für etwaige manuelle Korrekturen (in Abb. 1 auf Tier 1),
6. Kommentare, z. B. wenn ein Wortakzent falsch realisiert wurde (in Abb. 1 auf Tier 6).

In Abbildung 1 kann man beispielhaft die sechs Annotationsebenen für die Aufnahme des Satzes *Ich wüsste nicht, wie der schnellste Weg nach Polen ist.* eines Lernalters gut erkennen. Grundlage für die Repräsentation in Lautschrift war eine normative (kanonische) Transkription der vorzulesenden Wörter. Im nachfolgenden Schritt wurden die Wort- und Lautgrenzen automatisch durch so genanntes „forced alignment“ festgesetzt, einem Verfahren aus der automatischen Spracherkennung, bei der die angenommenen produzierten Laute zeitlichen Abschnitten des Sprachsignals zugeordnet werden (hier durch ein intern entwickeltes Programm). Im letzten Schritt wurden, falls notwendig, manuelle Korrekturen der Laut- und Wortgrenzen vorgenommen und Ersetzungen, Ausfälle und Einfügungen von Lauten annotiert. Die manuellen Korrekturen wurden durch mehrere studentische Hilfskräfte (mit entweder Französisch oder Deutsch als L1) durchgeführt, und zwar für ca. 80 Prozent der beiden L2-Sub-Korpora, was für ein



**Abb. 1:** Ein Beispielsatz dargestellt als Sprachsignal (Zeitsignal oben, Spektrogramm unten) mit den sechs Annotationsebenen (siehe Text für Details).

Korpus dieser Größenordnung als viel betrachtet werden kann. Dabei stand die Bearbeitung der ersten beiden Teile (Vorleseaufgabe und Nachsprechaufgabe) im Vordergrund. Es ist aber auch geplant, dass weitere Annotationen dem Korpus im Laufe der Zeit hinzugefügt werden, z. B. bez. der Pausen im vierten Teil (Vorlesegeschichte).

## 3 Anwendungsbeispiele

### 3.1 Messphonetische Studien der Produktion und Perzeption

/h/

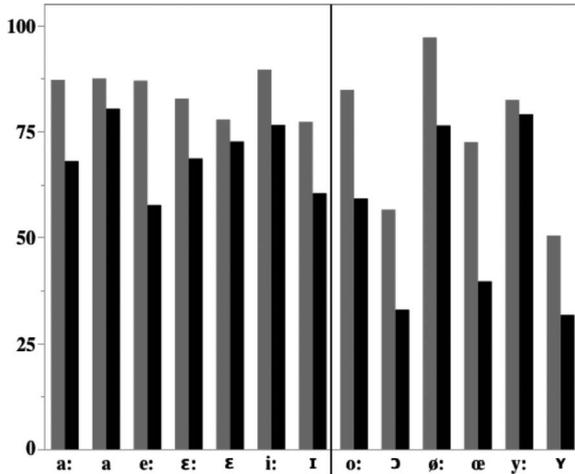
Ein für deutsche Hörer typisches Merkmal des französischen Akzents im Deutschen ist die fehlende Realisierung von /h/ in Wörtern wie *Hase* oder *Rathaus*. Minimalpaare wie *Hund* – *und* und *Haus* – *aus* zeigen, dass diese Interferenz auch zu Problemen bei der Verständigung führen kann und damit mehr als nur einen vielleicht lustig klingenden Akzent darstellt. Das Problem des fehlenden glottalen Frikativs, als der /h/ üblicherweise dargestellt wird, kann noch verstärkt werden, indem /h/ durch den Glottalverschluss, einen im Französischen ebenfalls nicht üblichen Konsonanten, ersetzt wird – und umgekehrt ein Glottalverschluss durch [h]. Im Korpus waren daher Realisierungen wie [ham aus haus ɔlts] für *ein Haus aus Holz* mit /h/-Aussprache für den Glottisverschluss [ʔ] nicht selten.

Es ist daher interessant, mehr darüber zu erfahren, ob Anfänger grundsätzlich alle /h/ löschen und inwieweit Fortgeschrittene die korrekte Realisierung von /h/ beherrschen. In Zimmerer & Trouvain (2015a) wurden im Korpus 34 Wörter analysiert, die mit /h/ beginnen. Es zeigte sich, dass Anfänger, wie vermutet, /h/ gänzlich ausfallen lassen, aber auch – und das ist das Neue – dass sie in allererster Linie den Glottisplosiv oder eine andere Art von Glottalisierung produzierten, also [ʔamʔaus] für *ein Haus*. Fortgeschrittene waren erfolgreicher, wenn es darum ging, /h/ so wie L1-Sprecher zu realisieren. Dabei wird im Deutschen /h/ je nach Stimmhaftigkeit des vorangehenden Kontextes als [h] oder [ɦ] produziert, also *das Haus* stimmlos mit [h], aber *ein Haus* stimmhaft mit [ɦ]. Dieser Befund nährt die Vermutung, dass L1- und fortgeschrittene L2-Sprecher bei [h] keine oder eine geringe Adduktion der Stimmlippen zeigen, L2-Anfänger hingegen einen hohen Grad der Stimmlippenadduktion, was zu Hyperartikulation führt. Unklar bleibt aber, durch welche Maßnahmen fortgeschrittene Lerner diese mutmaßliche Entspannung der Stimmlippenaktivität erworben haben und wie man Lerner auf ihrem Weg dabei unterstützen kann. Hierzu wären ergänzende instrumental-phonetische Experimente sinnvoll.

### Qualität und Quantität von Vokalen

Die Korpus-Daten können auch als Grundlage für Perzeptionstests dienen, wie in Zimmerer & Trouvain (2015b) gezeigt wird. In den Vorlese-Sätzen des Korpus versteckten sich unter anderem Minimalpaarwörter, die sich in Vokalquantität und -qualität unterscheiden, wie z. B. *Höhle* – *Hölle*, *schief* – *Schiff*, *Wüste* – *wüsste*. Beispiele für Wortpaare, die sich nur durch Quantität des Vokals, aber nicht durch dessen Qualität unterscheiden, sind *Täler* – *Teller* und *Kahn* – *kann*. Die betreffenden Wörter der Minimalpaare (n=11) der französischen Sprecher wurden ausgeschnitten und L1-Sprechern zur Unterscheidung vorgespielt („Höhle oder Hölle?“).

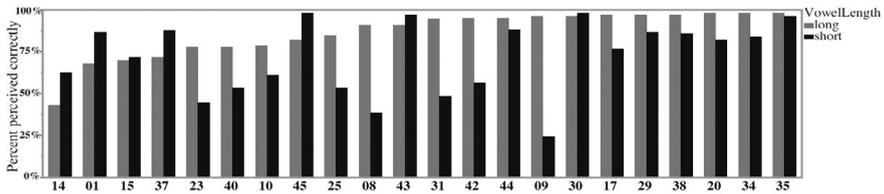
Ein Vergleich der Vokalinventare des Deutschen und des Französischen ließe erwarten, dass zum einen die im Französischen fehlenden hohen Vokale wie /ɪ/ oder /y/ und zum anderen die Quantitäts- und Qualitätsopposition Probleme bereiten. Einerseits gibt es die im Französischen unbekannte Unterscheidung in Kurz- und Langvokale nicht, andererseits unterscheiden sich die meisten Lang- und Kurzvokalpaare im Deutschen durch ihre Gespanntheit und werden daher durch ihre Vokalqualität in ungespannte (und kurze) und gespannte (lange) Vokale unterteilt. Die Ergebnisse in Abb. 2 zeigen, dass vor allem die gerundeten Vokale ganz erhebliche Fehlinterpretationen zur Folge hatten. Als Beispiel wurden zwei Drittel aller Realisierungen von *Pollen* von den L1-Hörern als *Polen*



**Abb. 2:** Prozent korrekt interpretierter Vokale durch L1-Hörer je nach Vokal und Sprachbeherrschungsgrad: Anfänger (grau) oder Fortgeschrittene (schwarz) (die vertikale Linie trennt die ungerundeten von den gerundeten Vokalen). Die Prozentwerte von deutschen L1-Sprechern erreichen 100 % oder nahezu 100 %.

wahrgenommen, obwohl [ɔ] im Französischen existiert. Unklar bleibt dabei aber, ob die Hörer mehr die Länge oder mehr die Qualität zur Entscheidung herangezogen haben. Zur Kontrolle wurden in einem Nachtest auch die entsprechenden Wörter von 40 L1-Sprechern untersucht, die, wie erwartet, nahezu zu 100 % als korrekt eingestuft wurden.

Ebenso erwartet war, dass die Fortgeschrittenen als Gruppe besser waren als die Anfängergruppe. Schaut man sich allerdings die individuellen Ergebnisse an, so wird deutlich, dass etliche Sprecher der Fortgeschrittenengruppe gravierende Defizite bei der Unterscheidung der Vokale aufweisen (s. Abb. 3). Hierbei lassen sich sowohl solche finden, die in erster Linie Probleme mit den gespannten Langvokalen haben (z. B. Sprecher 14), als auch solche mit Problemen bei den ungespannten Kurzvokalen (z. B. Sprecher 08, 09). Es gibt aber auch L2-Sprecher, die sehr gut bewertet werden, nämlich auf L1-Niveau (z. B. Sprecher 30, 35). Interessant sind Realisierungen von Sätzen wie *Im Frühling fliegen Polen (sic!) durch die Luft*, die nahezu frei von fremdem Akzent sind, aber klar eine phonologische Interferenz aufweisen. Die Analyse des Hörtests zeigt für jeden individuellen Lerner, welche Schwierigkeiten bei den Vokalen vorhanden sind und dass das allgemeine Niveau der Sprachbeherrschung nicht besonders zuverlässig mit den hier gezeigten Leistungen bei der Vokalproduktion korreliert.

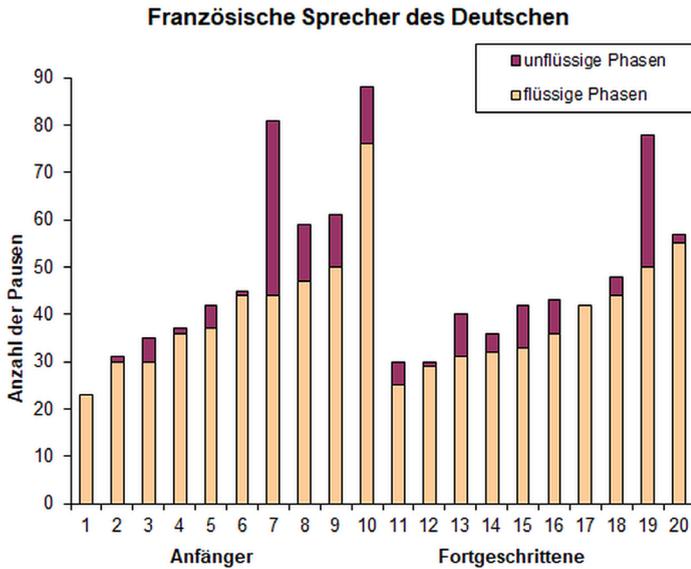


**Abb. 3:** Korrekte Zuweisung der Minimalpaarwörter („Höhle oder Hölle?“) in Prozent (y-Achse) für jeden der fortgeschrittenen Lerner (Sprecher-ID auf der x-Achse), geordnet nach den Werten für die gespannten Langvokale (grau) und die ungespannten Kurzvokale (schwarz).

### Sprechtempo und Redefluss

Sprechtempo und Redefluss zeigen immer wieder starke Korrelationen in Bezug der Niveaustufe der L2-Lerner, sei es im Vergleich zu ihrer eigenen L1, aber auch im Vergleich zu L1-Sprecher der Zielsprache. Neben dem Sprechtempo, das sich massiv in der Nutzung von Pausen manifestiert, können Pausen auch eine große Rolle als Disfluenzen spielen, sei es als eine überlange Pause, die als Häsitiation interpretiert wird, oder als ein Mittel zum Editieren in einer Reparatur (z. B. Bosker et al. 2015, Trouvain & Braun 2020). Die Vorlesegeschichte *Die drei kleinen Schweinchen* (1–2 Minuten Sprechzeit) eignet sich sehr gut, um Vergleiche zu Sprechtempo und Redefluss mit dem *identischen* Material anzustellen – im Gegensatz zu Spontansprache, in der sich über eine Minute hinweg das Material über Produktionen hinweg nicht identisch bleibt.

Trouvain, Fauth & Möbius (2016) untersuchten die Pausen von französischen Sprechern (10 Anfänger und 10 Fortgeschrittene) aus dem IFCASL-Korpus und fanden in deren L2-Daten wie zu erwarten mehr Pausen, längere Pausen, sowie mehr Disfluenzen als in den französischen L1-Daten und den deutschen L1-Daten (vgl. Abb. 4). Dabei war auffällig, dass die Pausen in den unflüssigen Phasen im Vergleich zu denen aus den flüssigen Phasen 1) kürzer waren, 2) nie ein Atemgeräusch und 3) nie eine Füllpartikel wie *äh* oder *ähm* enthielten. Der letztgenannte Befund steht sehr wahrscheinlich in Zusammenhang mit unterschiedlichen Planungsprozessen bei Lese- und Spontansprache und sollte mitbeachtet werden, wenn Disfluenzen und Sprechfluss von L2-Produktionen betrachtet und bewertet werden (vgl. Hilton 2009, Bosker et al. 2015, Belz & Odebrecht, in diesem Heft).



**Abb. 4:** Anzahl von Pausen (in flüssigen und unflüssigen Phasen) für jeden der 10 L2-Sprecher der Anfänger (1–10) und der Fortgeschrittenen (11–20).

### 3.2 Lehre von DaF und in der Ausbildung von DaF-Lehrkräften

Der Einsatz von Daten aus einem phonLK für die Lehre in Deutsch als Fremdsprache (DaF) und in der Ausbildung von DaF-Lehrkräften bietet vielfältige Möglichkeiten. Die hier exemplarisch dargestellten Ideen mögen als Optionen und auch als Anregung zur Nutzung dienen.

Der Datenreichtum eines phonLK kann bei der Erstellung von *Übungsaufgaben* in Bezug auf die perzeptuelle Sensibilisierung mittels der Audiodateien genutzt wird. Audiobeispiele in fremdsprachlichem Lehrmaterial enthalten in aller Regel ausschließlich Aufnahmen von L1-Sprechern der Zielsprache. Zur „auditiven Anschauung“ ließen sich bei Daten eines phonLK wie dem IFCASL-Korpus, aber auch anderer gLK, derart nutzen, dass für dieselben Beispiele auch Aufnahmen von L2-Sprechern herangezogen werden können, um z. B. Unterschiede in ausgewählten Details zu besprechen. Die Aufgabe könnte zunächst darin bestehen Aufnahmen desselben Satzes von vier verschiedenen Sprechern (zwei L1-Sprecher, zwei L2-Sprecher, davon einer ein Anfänger) vorzulegen und zu fragen, wer der vier L1-Sprecher ist und wer L2-Sprecher auf Anfängerniveau. Die Folgeauf-

gabe könnte im auditiven Analysieren eines Details wie der Aspiration von Plosiven bestehen, indem auch immer wieder versucht wird, die vier Aufnahmen zu vergleichen. Individuelle Lerner könnten auch individuelle Kombinationen von Sprechern erhalten. Idealerweise würde das vergleichende Hören und Analysieren nicht nur mit auditiven, sondern auch mit visuellen Illustrationen ergänzt werden.

Durch das Vorhandensein von Aufnahmen vieler Sprecher kann auch ein so genanntes *high variability pronunciation training* (HVPT) stattfinden, bei dem der Lerner mit vielen Modellsprechern statt nur einem Modellsprecher konfrontiert wird. Eine Metaanalyse zu HVPT von Thomson (2018) legt nahe, dass dieses bei Sprachvermittlern weitgehend unbekannte Verfahren sehr effektiv ist. Nicht nur phonLK, sondern gesprochene Korpora im Allgemeinen bilden eine hervorragende Grundlage für *high variability training*, wenn dabei auch die Zielsprache durch L1-Sprecher abgedeckt wird. Bei einem bidirektionalem phonLK wie dem IFCASL-Korpus oder dem Korpus von García Lecumberri et al. (2016) ist dies der Fall.

Für die Ausbildung von DaF-Lehrkräften kann ein phonLK helfen, die *Analyse von Ausspracheproblemen* einzuüben. Hörbare Lernerdaten, die direkt miteinander vergleichbar sind, können als sinnvolle Ergänzung zur eher theoretischen Besprechung möglicher Interferenzen auf phonetischer und phonologischer Ebene dazu herangezogen werden. Dabei können auch Aspekte zu Kompetenzbeschreibungen konkret angeschaut und angehört werden, wie sie im GERS genannt werden, um „die Terminologie der Deskriptoren zu den Niveaustufen, eine inkonsistente Bezugnahme auf segmentale und suprasegmentale Aspekte der Aussprache sowie die schlechte Trennbarkeit von phonetisch-phonologischen Aspekten und solchen wie Sprechflüssigkeit und grammatische Richtigkeit bei der Bewertung der Verständlichkeit“ zu diskutieren (Dahmen 2019: 195). Von größerem Interesse dürfte dabei sein, wie man Aussprachefehler und Ausspracheabweichungen interpretiert und woher diese Interferenzerscheinungen rühren, wie z. B. unzureichende Kenntnis und Umsetzung von Orthographie-zu-Aussprache-Regeln (bei reiner Lesesprache), vermutete Unkenntnis bestimmter Wörter, inkorrekte phonologische Kategorisierung (z. B. bei /ç/) oder der Umgang mit Kognaten wie international benutzte geographische Namen. Der Einsatz weiterer gLK würde aber auch Vergleiche zwischen lesesprachlichen und spontan-sprachlichen Leistungen von Lernern erlauben.

### 3.3 Entwicklung von computergestütztem Aussprachetraining

Je schneller computergestützte Verfahren beim Sprachenlernen an Bedeutung gewinnen, desto wichtiger wird es, die Verzahnung von technologischen Möglichkeiten, didaktischen Anforderungen und linguistisch-phonetischem Wissen

effizient *und* effektiv zu kontrollieren. Für den phonetischen und phonologischen Bereich finden bei Programmen des computerassistierten Aussprachetrainings (computer-assisted pronunciation training = CAPT) Entwicklungen statt, die sowohl Lese- als auch Spontansprache berücksichtigen (O'Brien et al. 2019, Zechner & Evanini 2019), bei der aber eine kritisch-konstruktive Sicht der Linguisten auf die technischen und pädagogischen Möglichkeiten notwendig bleibt (Neri et al. 2008). Ein phonLK bietet für die Entwicklung von CAPT-Anwendungen ein sehr gutes Testfeld für phonetisch-didaktische Fragestellungen.

Wortakzent spielt eine wichtige Rolle im Deutschen (Hirschfeld & Reinke 2018) und stellt für Deutsch-Lerner häufig eine große Hürde dar – unabhängig davon, ob dies den Lernern bewusst ist oder nicht. Für Lerner mit Französisch als L1 gilt dieser Umstand ganz besonders, da Wortakzent im Französischen keine Rolle spielt: es wird stets die letzte Silbe betont (außer wenn diese mit Schwa produziert werden sollte). Hörexperimente mit französischen L1-Sprechern haben gezeigt, dass diese oftmals große Probleme haben sich die betonten Silben als lexikalische Information abzuspeichern, so dass sogar von „stress-deafness“ (also „Betonungstaubheit“) gesprochen wird (Dupoux et al. 2008). Solche Lerner könnten also von einem computer-assistiertem Aussprachetraining profitieren, welches individualisiertes korrekatives Feedback liefert. Voraussetzung für ein sinnvolles Feedback ist die verlässliche automatische Detektion solcher Abweichungen in der Sprachproduktion der individuellen L2-Lerner. Vakil & Trouvain (2015) präsentieren einen Ansatz, bei dem maschinelles Lernen zur Klassifikation von Akzentzuweisungsfehlern zum Einsatz kommt. Grundlage für die Trainings- und die Überprüfungsdaten stammen dabei aus dem IFCASL-Korpus.

Ein anderes Beispiel zur Nutzung des IFCASL-Korpus zu CAPT-Anwendungen ist Feedback zur Verbesserung der Qualität und Quantität der Vokale im Deutschen. Für die meisten Deutsch-Lerner stellen das vergleichsweise große Inventar an Vokalen eine schwer meisterbare Schwierigkeit dar. Neben den Kontrasten in der Vokalqualität gesellt sich noch die Vokalquantität. Letzteres ist auch für Französisch-Sprecher ein Problem wie oben in Sektion 3 gezeigt worden ist. Dieses Problem betrifft nicht nur die Produktion, sondern auch die Perzeption. Carroll et al. (2015) stellen ein visuelles Feedbacktool vor, bei dem die Lerner die Aufgabe haben, sich Wörter aus Minimalpaaren anzuhören und danach zu wiederholen. Aus der Aufnahme des Lerners wird das akustische Korrelat des betonten Vokals erkannt und seine Formantenstruktur (F1 und F2) auf einer Art Zielscheibe abgebildet. Zusätzlich wird auch die relative Dauer des Vokals dargestellt. Die eigenen Produktionen können dabei stets mit solchen von L1-Sprechern visuell und auditiv verglichen werden. Ziel ist es, dass sich durch die Wiederholungen und die Vergleiche eine Sensibilisierung für die notwendigen Kontraste aufbaut und die artikulatorischen Ziele näher an L1-Vokalproduktionen erreicht werden.

Die Nutzung vieler Modell-Sprecher für Minimalpaarwörter aus dem Korpus kann hierbei wertvolle Dienste leisten.

Eine ähnliche Herangehensweise verfolgen Jügler & Zimmerer (2017). Hierbei steht nicht eine sichtbar gemachte Annäherung an das artikulatorische Ziel im Vordergrund, sondern eine kategorische Rückmeldung, ob Vokalqualität und -dauer in einem akzeptablen Bereich liegen. Visualisiert wird dies durch einen Frosch, der beim Hüpfen (ausgelöst durch das vom Lerner artikulierte Wort) die nahe gelegene Seerose erreicht und sicher dort landet – oder daneben ins Wasser fällt.

Großes Potenzial für Übungen in CAPT liegt auch in satzprosodischen Bereichen wie dem Sprechtempo der Zielsprache, das für viele Lerner ein großes Hindernis darstellt. Auf der Perzeptionsseite wäre es wünschenswert, wenn ein Lerner sich Texte so schnell bzw. langsam vorspielen lassen kann wie individuell gewünscht. Dies kann durch Auswahl verschieden schneller L1-Sprecher geschehen, aber auch unter Zuhilfenahme geeigneter Signalmanipulation. Auf der Produktionsseite könnten Mitsprechübungen angeboten werden, bei denen der Lerner einen gegebenen Text zum einen sieht und zum anderen von einem L1-Sprecher hört mit der Aufgabe, den Text laut und synchron mit dem Modellsprecher mitzulesen. Dabei sollten die gerade vom L1-Sprecher produzierten Wörter visuell markiert werden. Auch hier sollte der Lerner das Tempo individuell anpassen können.

## 4 Beschränkungen

### 4.1 Fehlende Spontansprache

Rein vorgelesenes Material hat sowohl Vor- als auch Nachteile für verschiedene Forschungsfragen. Im Vergleich zu Spontansprache hat Lesesprache den Vorteil, dass man alle Segmente der jeweiligen Sprache systematisch elizitieren kann. Insbesondere lassen sich gezielt Minimalpaare, Kognaten, Lehnwörter, Zahlen und Abkürzungen einsetzen. Zudem ist – im Gegensatz zu Spontansprache – das gesamte Material direkt vergleichbar. Ferner ist Lesematerial unerlässlich, um Probleme bei der Graphem-Phonem-Beziehung in der L2 zu untersuchen, was vor allem bei Lernern auf Anfänger-Niveau vorkommen wird. Leseaufnahmen können sich schließlich auch dazu eignen, um die visuell-graphische Modalität mit der Modalität des Zuhörens zu kombinieren, z. B. durch Einbeziehung eines Modellsprechers, so dass eine kombinierte Lese- und Nachsprechaufgabe entsteht (Detey et al. 2010 und Trouvain et al. 2016).

Klarerweise hat die Beschränkung auf Lesematerial aber auch Nachteile. Der wichtigste ist sicherlich, dass wir im Alltag mehr spontan sprechen als lesen. Beim Sprechen und Hören von Fremdsprachen ist der spontane Anteil an Produktion und Rezeption von Sprache oftmals reduziert im Vergleich zur Muttersprache, da beim Lernen in der Regel schriftlicher Input dominiert. Weitere „authentische“ Formen von Lesesprache begegnet allen Sprachteilnehmern durch Textsorten, die man als „written to be spoken“ charakterisieren könnte. Dazu gehören verschiedene Textsorten, die in diversen hör-basierten Medien oftmals durch professionelle Sprecher aufgenommen werden, z. B. Nachrichten, Reportagen, Werbespots im Radio und im Fernsehen, aber auch Sätze und Texte in Sprachlehrwerken. Eine besondere und sehr wichtige Form von schriftbasierter Sprache im Lernkontext stellen Liedtexte dar, da sie für Lerner ein nicht zu unterschätzendes Modell zum Mithören und zum Mitartikulieren darstellen.

Trotz der alltäglichen Bedeutung der Spontansprache ergibt sich forschungsökonomisch aber auch das Problem, dass entsprechende Aufnahmen und Analysen sehr aufwändig sind. Aus diesem Grund beschränkt sich das IFCASL-Korpus (Trouvain et al. 2016) auf Leseaussprache. Phonetische Lernerkorpora müssen sich allerdings nicht ausschließlich auf Lesesprache beziehen. Ein gutes Beispiel für ein bidirektionales Korpus mit spontansprachlichen Dialogen ist bei García Lecumberri, Cooke & Wester (2017) beschrieben. Die Aufnahmen mit Sprechern des Spanischen und des Englischen wurden sowohl in ihrer jeweiligen L1 und in ihrer L2 durchgeführt. Um eine Vergleichbarkeit zu gewährleisten wurden so genannte Diapix-Aufgaben durchgeführt (siehe auch Belz & Odebrecht, in diesem Heft), bei denen die Sprecherpaare Unterschiede in „Wuselbildern“ gemeinsam herausfinden mussten (jeder der beiden Sprecher hatte leicht unterschiedliche Szenerien). Die Spontansprache hat hier den Nachteil, dass nicht exakt das gleiche Material für alle Sprecher vorliegt, was auch einen größeren Aufwand bei der (automatischen) Annotation der segmentellen Struktur zur Folge hat.

## 4.2 Instrumentelle Erweiterungen

Auch wenn sich die Primärdaten von phonLK bislang auf Audiodateien beschränken, so ist auch denkbar, dass andere Arten phonetischer Daten mit aufgenommen werden, denn experimentelle phonetische Forschung, welche die Sprechproduktion von Lernern untersucht, ist nicht auf Audiodaten beschränkt. Beispiele hierfür sind die Studien mit Ultraschallaufnahmen von Gick et al. (2008) und Kocjančič Antolík et al. (2019) sowie eine Studie von Wieling et al. (2015), die Lerner-Aufnahmen mit elektromagnetischer Artikulographie (EMA) untersuchen. Im dem hier vorliegenden Aufsatz werden als primär erfasste Daten aber aus-

schließlich Mikrophonesignale in Form von Audiodateien behandelt. Denkt man aber einen Schritt weiter, so werden in Zukunft verstärkt Forschungsdaten über Mikrophonesignale hinausgehen, die in experimentellen Forschungsfragen, aber auch in technologischen Anwendungen ihren Niederschlag finden (z. B. visuelle Animationen von Zunge und anderen Artikulatoren sowie der Phonation und der Sprechatmung in der Zielsprache).

### 4.3 Datenmanagement

Mehr Sensibilisierung in der linguistischen Forschung für Aspekte des Datenmanagements ist der Erfahrung nach im stärkeren Maße notwendig. Fragen des Datenmanagements bei Korpora sind sehr wichtig und keinesfalls „lästiges Beiwerk“, weswegen an dieser Stelle gesondert hingewiesen wird in Ergänzung zu Hirschmann & Schmidt (in diesem Heft). Allerdings können die dargelegten Punkte hier nur kurz angerissen werden.

Datensammlungen wie sprachliche Korpora benötigen eine besondere Art und einen besonderen Umfang der Pflege. Im Allgemeinen gilt, dass Daten aus Korpora nach den so genannten FAIR-Prinzipien zur Verfügung gestellt werden sollten. Die Daten sollten für berechnigte Interessierte folgende Merkmale aufweisen: gut auffindbar (*findable*), problemlos zugänglich (*accessible*), in puncto Formaten gut dokumentiert (*interoperable*) und für andere Forschungsvorhaben wiederverwendbar (*reusable*).

Gelten die FAIR-Prinzipien generell für alle Datensammlungen, so müssen bei gLK an manchen Stellen Besonderheiten beachtet werden. Die Wiedererkennbarkeit individueller Sprecher ist bei gesprochenen Korpora um ein Vielfaches höher als bei rein schriftlichen Korpora. Auf diesen Umstand sollte hingewiesen werden, wenn die aufgenommenen Sprecher der Nutzung zur Forschung durch die Datensammler und auch durch Drittnutzer schriftlich zustimmen. Die potenzielle Weitergabe der Daten an Dritte stellt besondere Herausforderungen an den Zustimmungssakt, weil zum einen sichergestellt werden soll, dass die Sprecher als Datenspende über die gesamte Breite der Nutzung ihrer Daten Bescheid wissen sollen, und zum anderen die Datensammler sich von den Folgenutzern versichern lassen möchten, in welchem Rahmen die Daten verwendet und bearbeitet werden dürfen. Eine ungeklärte Frage betrifft den Datenschutz bei Audiodateien. Auch wenn es bei verwendeten Sprecherkürzeln zu einer echten Anonymisierung und nicht nur zu einer Pseudonymisierung gekommen sein sollte, so scheint eine Anonymisierung auf Grundlage der individuellen Stimme nicht zu 100 % gewährleistet. Dies kann weit reichende Folgen für die Verbreitung und Nutzung der Daten haben, wenn aus Gründen der potenziellen Erkennbarkeit der

Individuen der Zugang zu diesen Daten stark eingeschränkt wird. Zentral dabei ist eine frühzeitige und ausführliche Auseinandersetzung mit der Aufklärung der Datenspender, die im Falle von Lernern auch nicht erwachsene Personen sein können, um ihre Zustimmung für die weitreichende Nutzung zu erhalten. Ein weiteres Spannungsfeld, auf das an dieser Stelle verwiesen wird, ist dass die Repositorienlandschaft zuweilen nicht darauf eingestellt ist, versionierte Korpora zu verwalten, also sinnvolle Erweiterungen (auch anderer Forschungsgruppen) zur Verfügung zu stellen.

Bezüglich von Formaten der Daten gesprochener Sprache stellen sich zwei fundamentale Herausforderungen. Zum einen betrifft dies das Speicherformat der Signaldateien: es besteht keine Garantie, dass im wav-Format abgespeicherte Audiodateien auch noch in 10 oder 20 Jahren genauso verwertbar sind. Sie müssten also notfalls konvertiert werden. Zum anderen betrifft es das Format der Annotationsdateien. Auch wenn in den letzten Jahrzehnten der Speech-Editor Praat (Boersma & Weenink 2016) eine Art Standard zur phonetisch-akustischen Analyse und zur Annotation (als so genannte TextGrids) betrachtet werden kann, so gibt es auch hier keine Garantie, dass die Annotationsdateien in 20 Jahren noch anwendbar sind. Es sollte auch noch darauf hingewiesen werden, dass es Nutzern bewusst sein sollte, dass die Speicherung gesprochener Daten z. B. im wav-Format ein Vielfaches an Speicherplatz zu Textdateien erfordert.

Metadaten spielen generell eine zentrale Rolle im Forschungsdatenmanagement. Für gesprochene Daten sind hier entscheidend, dass möglichst viele Metadaten zu den Sprechern aufgenommen werden, also Alter, Geschlecht, regionale Herkunft, höchster Bildungsabschluss, Dialektbenutzung, Erst-, Zweit- und Drittsprachen, Beeinträchtigungen der Sprech-, Denk-, Hörfähigkeiten. Im Fremdsprachenlernbereich sind natürlich Informationen zur Einschätzung des Niveaus, Jahre der Ausbildung sowie die Dauer relevanter Auslandsaufenthalte von Interesse.

All die genannten Erfordernisse stellen viele Herausforderungen an das Datenmanagement gesprochener Lernerkorpora im Allgemeinen dar. Idealerweise werden die genannten Aspekte vor Beginn der Datenerhebung eindeutig geklärt, genau genommen gehören sie zur Konzeptionierung des Projekts. Eine Unterbrechung bzw. eine Nichtbeachtung dieser Aspekte wird üblicherweise zur Folge haben, dass das entsprechende Korpus nicht weiter genutzt werden kann und somit für Folgeforschung nutzlos bleibt und die durch die Datensammler durchgeführten Studien nicht reproduzierbar sind.

## 4.4 Einsatz für Lehrkräfte

Es wäre naiv anzunehmen, dass Lehrkräfte Daten aus gLK, speziell phonLK, sofort nach Bedarf einsetzen können. Weder haben die Datensammler die Lehrkräfte und ihre genauen Bedürfnisse bei der Aufbereitung der Daten im Blick, noch sind die Lehrkräfte darin geschult in einem Lernerkorpus (auf die Schnelle) genau das zu finden, was sie für bestimmte Aufgaben und Übungen suchen. Hierzu bräuchte es eine Vermittlungsglied, das es in der optimalen Ausprägung noch nicht zu geben scheint. Daher bestehen bislang die Vermittlungsversuche zwischen Lernerkorpus einerseits und Lehrkräften als potenziellen Nutzern andererseits, dass interessierte Studenten sich mit Teilen des Lernerkorpus auseinandersetzen, um damit selbst Unterrichtsentwürfe zu entwickeln (z. B. Brunstein 2016).

Es bleibt momentan ein Desiderat, eine funktionierende Schnittstelle zwischen Korpus und seinem Einsatz im Unterricht zu haben. Schritte dazu müssen zuerst in der studentischen Ausbildung der Lehramtsstudiengänge erfolgen, aber auch im Herausfinden der Bedürfnisse auf der Seite der Lehrkräfte.

## 5 Zusammenfassung und Ausblick

Der vorliegende Aufsatz hat das Potenzial phonetischer Lernerkorpora beschrieben. So stellen phonLK für die *Grundlagenforschung* eine sehr wertvolle Ressource für Forschungsdaten im L2- und im L1-Bereich dar. Auch allgemein sprachvergleichend können phonLK eingesetzt werden, zum Beispiel im hier beschriebenen Sprachenpaar Deutsch-Französisch, genauso wie die Analyse individueller Lerner auf verschiedenen Niveaustufen. Somit können phonLK als eigenständige Datengrundlage dienen, aber auch als Ergänzung für experimentelle Studien, die oftmals nicht an die hier vorhandene Datenmenge heranreichen.

Im Gegensatz zu *stummen* Lernerkorpora und *sprechenden* Lernerkorpora haben phonLK den Vorteil, dass die Audiosignale nicht nur hörbar sind, sondern die Annotationen bis auf Satz-, Wort- und Segmentebene auch zeitlich aligniert vorhanden sind. Dies schafft enorme Vorteile für die auditive Analyse phonetischer und phonologischer Fragen und erst die Voraussetzung für eine sinnvolle akustische Analyse solcher Fragen. Somit stellen phonLK auch eine gute Ergänzung für die *DaF-Forschung* dar, deren Fokus in aller Regel auf kontrastiven Analysen liegt.

Im Bereich der *DaF-Lehre* können phonLK große Potenziale bergen, um Übungen zu erstellen und diese auditiv mit vielen statt nur einem Modellsprecher

zu illustrieren. Ein Umstand, der auch für *computergestütztes Aussprachetraining* wichtig ist, da bei phonLK sowohl Daten für das Training als auch zur Demonstration in größerer Menge vorhanden sind.

PhonLK weisen neben Potenzialen auch Beschränkungen auf. Sowohl Lesesprache als auch Spontansprache haben ihre Vorteile, die man im Rahmen der Fremdsprachenvermittlung nutzen kann. Im hier beschriebenen Korpus ist ausschließlich Lesesprache vorhanden, in anderen phonLK ausschließlich Spontansprache.

Das Datenmanagement spielt für sprachliche Korpora generell eine nicht zu unterschätzende Rolle, da die Daten für relevante Gruppen nutzbar und zugänglich gemacht und auch zugänglich gehalten werden müssen. Dies schließt rechtliche, technische, aber auch didaktische Fragen mit ein. Was die letztendliche Nutzbarkeit für ganz verschiedenen potenziellen Benutzergruppen anbelangt, gibt es bislang noch einen Forschungsbedarf.

**Danksagung:** Der Aufsatz hat durch die gegenseitige Begutachtung durch die Kolleginnen und Kollegen der anderen Beiträge sehr profitiert, genauso wie durch die umsichtigen und wertvollen Kommentare der Gastherausgeberin, aber auch durch die Anmerkungen der Herausgeber/innen dieser Zeitschrift, wofür ich sehr danken möchte.

## Literatur

- Ballier, Nicolas & Martin, Philippe. 2015. Speech annotation of learner corpora. In S. Granger, G. Gilquin, & F. Meunier (Hg.), *The Cambridge Handbook of Learner Corpus Research*. Cambridge: Cambridge University Press, 107–134.
- Belz, Malte, Sauer, Simon, Lüdeling, Anke & Mooshammer, Christine. 2017. Fluently disfluent? Pauses and repairs of advanced learners and native speakers of German. *International Journal of Learner Corpus Research* 3(2). Special Issue on Phonetic Learner Corpora, 118–148. <http://hu-berlin.de/bematac>
- Boersma, Paul & Weenink, David. 2016. *Praat: doing phonetics by computer*. Version 6.0.19, <http://www.praat.org/>.
- Bosker, Hans R., Pinget, Anne F., Quené, Hugo, Sanders, Ted & De Jong, Nivja. 2015. What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing* 30(2), pp. 159–175.
- Brunstein, S. 2016. Der Erwerb der L2-Intonation als Herausforderung an die Fremdsprachendidaktik mit einer empirischen Studie zur französischen Intonation im L1-Deutsch Kontext. Abschlussarbeit. Universität Stuttgart.
- Carroll, Patrick, Trouvain, Jürgen, & Zimmerer, Frank. 2015. A visual feedback tool for German vowel production. Proc. 26. Konferenz Elektronische Sprachsignalverarbeitung (ESSV), Eichstätt, 150–157.

- CLEL – Centre for English Corpus Linguistics (17.03.2021): Learner Corpora around the World. Louvain-la-Neuve: Université catholique de Louvain. <https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-the-world.html>
- Chun, Dorothy & Levis, John M. 2020. Prosody in L2 teaching: methodologies and effectiveness. In: Gussenhoven, C. & Chen, A. (Hg.). *The Oxford Handbook of Language Prosody*. Oxford: Oxford University Press, 619–630.
- Dahmen, Silvia. 2019. Die Skalen zu Phonetik/Phonologie im GeR und seinem Begleitband. *Deutsch als Fremdsprache* 4/2019, pp. 195–204.
- Draxler, Christoph. 2008. *Korpusbasierte Sprachverarbeitung. Eine Einführung*. Tübingen: Gunter Narr Verlag.
- Dupoux, Emmanuel, Sebastián-Gallés, Núria, Navarrete, Eduardo, Peperkamp, Sharon. 2008. Persistent stress ‚deafness‘: the case of French learners of Spanish. *Cognition* 106(2), 682–706.
- Durand, J., Gut, U., Kristoffersen, G. (Hg.) 2014. *Oxford Handbook of Corpus Phonology*. Oxford: Oxford University Press.
- García Lecumberri, María L., Cooke, Martin & Wester, Mirjam. 2017. A bi-directional task-based corpus of learners' conversational speech. *International Journal of Learner Corpus Research* 3(2), 175–195.
- Gut, Ulrike. 2014. Corpus phonology and second language acquisition. In: Durand, J., Gut, U., Kristoffersen, G. (Hg.) *The Oxford Handbook of Corpus Phonology*. Oxford: Oxford University Press, 286–301.
- Gick, Bryan, Bernhardt, Barbara M., Bacsfalvi, Penelope, & Wilson, Ian. 2008. Ultrasound imaging applications in second language acquisition. In: J. Hansen & M. Zampini (Hg.) *Phonology and Second Language Acquisition*, Amsterdam: Benjamins, 309–322.
- Gilquin, Gaëtanelle, De Cock, Sylvie & Granger, Sylviane. 2010. *Louvain International Database of Spoken English Interlanguage*, Louvain-la-Neuve: Presses universitaires de Louvain. <https://uclouvain.be/en/research-institutes/ilc/cecl/lindsei.html>
- Hilton, Heather. 2009. Annotation and analyses of temporal aspects of spoken fluency. *CALICO Journal* 26(3), pp. 644–661.
- Hirschfeld, Ursula & Reinke, Kerstin. 2018. *Phonetik im Fach Deutsch als Fremd- und Zweitsprache. 2., neu bearbeitete Auflage*. Berlin: Erich Schmidt Verlag.
- Jügler, Jeanin & Zimmerer, Frank. 2017. A computer-assisted vowel trainer for the German Language. *Proc. 28. Konferenz Elektronische Sprachsignalverarbeitung (ESSV)*, Saarbrücken, 340–345.
- Kocjančič Antolík, Tanja, Pillot-Loiseau, Claire & Kamiyama, Takeki. 2019. The effectiveness of real-time ultrasound visual feedback on tongue movements in L2 pronunciation training – Japanese learners' progress on the French vowel contrast /y/-/u/. *Journal of Second Language Pronunciation* 5(1), 72–97.
- Neri, Amra, Cucchiari, Catia, Strik, Helmer & Boves, Lou. 2008. The pedagogy-technology interface in computer assisted pronunciation training. *Computer assisted language learning* 15 (5), 441–467.
- O'Brien, Mary G., Derwing, Tracey M., Cucchiari, Catia, Hardison, Debra M., Mixdorff, Hansjörg, Thomson, Ron I., Strik, Helmer, Levis, John M., Munro, Murray J., Foote, Jennifer A., Levis, Greta M. 2018. Directions for the future of technology in pronunciation research and teaching. *Journal of Second Language Pronunciation* 4(2), 182–207.

- Thomson, Ron I. 2018. High Variability [Pronunciation] Training (HVPT) – A proven technique about which every language teacher and learner ought to know. *Journal of Second Language Pronunciation* 4(2), 208–231.
- Trouvain, Jürgen, Bonneau, Anne, Colotte, Vincent, Fauth, Camille, Fohr, Dominique, Jouviet, Denis, Jügler, Jeanin, Laprie, Yves, Mella, Odile, Möbius, Bernd & Zimmerer, Frank. 2016. The IFCASL corpus of French and German non-native and native read speech. *Proceedings of the 9th Language Resources and Evaluation Conference (LREC), Portorož*, 1333–1338.
- Trouvain, Jürgen & Zimmerer, Frank. 2016. Phonetische Lernerkorpora und ihr Nutzen im DaF-Bereich. *Deutsch als Fremdsprache* 53(4), 204–212.
- Trouvain, Jürgen, Fauth, Camille & Möbius, Bernd. 2016. Breath and non-breath pauses in fluent and disfluent phases of German and French L1 and L2 read speech. *Proc. Speech Prosody (SP8)*, Boston, pp. 31–35.
- Trouvain, Jürgen, Zimmerer, Frank, Möbius, Bernd, Gósy, Mária & Bonneau, Anne. 2017. Segmental, prosodic and fluency features in phonetic learner corpora – Introduction to the special issue. *International Journal of Learner Corpus Research* 3(2), 105–118.
- Trouvain, Jürgen & Braun, Bettina. 2020. Sentence prosody in a second language. In: Gussenhoven, C. & Chen, A. (Hg.). *The Oxford Handbook of Language Prosody*. Oxford: Oxford University Press, 605–618.
- Vakil, Anjana & Trouvain, Jürgen. 2015. Automatic classification of lexical stress errors for German CAPT. *Workshop on Speech and Language Technology in Education (SLaTE 2015)*, Leipzig, 47–52.
- Wieling, Martijn, Veenstra, Pauline, Adank, Patti, Weber, Andrea, & Tiede, Mark. 2015. Comparing L1 and L2 speakers using articulography. *Proc. International Congress of Phonetic Sciences, Glasgow*, Paper 551.
- Zechner, Klaus & Evanini, Keelan. 2019. *Automated Speaking Assessment: Using Language Technologies to Score Spontaneous Speech*. New York, NY: Routledge.
- Zimmerer, Frank & Trouvain, Jürgen. 2015a. Productions of /h/ in German: French vs. German speakers. *Proc. Interspeech, Dresden*, 1922–1926.
- Zimmerer, Frank & Trouvain, Jürgen. 2015b. Perception of French speakers' German vowels. *Proc. Interspeech, Dresden*, 1720–1724.