

Flüssigkeit und Korrektheit – Flüssigkeit oder Korrektheit?

—
Fluidité et correction – fluidité ou correction ?

—
Fluenza e correttezza – fluenza o correttezza?

—
Fluency and Accuracy – Fluency or Accuracy?

Peter Lenz, Malgorzata Barras, Fabienne Manz

2017

Bericht des Wissenschaftlichen Kompetenzzentrums für Mehrsprachigkeit
Rapport du Centre scientifique de compétence sur le plurilinguisme
Rapporto del Centro scientifico di competenza per il plurilinguismo
Report of the Research Centre on Multilingualism

Herausgeber | Publié par

Institut für Mehrsprachigkeit
www.institut-mehrsprachigkeit.ch

—
Institut de plurilinguisme

www.institut-plurilinguisme.ch

AutorInnen | Auteurs

Peter Lenz, Malgorzata Barras, Fabienne Manz

Übersetzung | Traduction

Mary Carozza, Roxane Barclay, Sandor Marazza

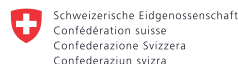
Das vorliegende Projekt wurde im Rahmen des Arbeitsprogramms 2012–2014 des Wissenschaftlichen Kompetenzzentrums für Mehrsprachigkeit durchgeführt. Für den Inhalt dieser Veröffentlichung sind die AutorInnen verantwortlich.

Le projet dont il est question a été réalisé dans le cadre du programme de travail 2012–2014 du Centre scientifique de compétence sur le plurilinguisme. La responsabilité du contenu de la présente publication incombe à ses auteurs.

Freiburg | Fribourg, 2017

Layout

Billy Ben, Graphic Design Studio



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Departement des Innern EDI
Département fédéral de l'intérieur DFI
Dipartimento federale dell'interno DFI
Departament federal da l'intern DFI
Bundesamt für Kultur BAK
Office fédéral de la culture OFC
Ufficio federale della cultura UFC
Uffizi federal da cultura UFC



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Eidgenössisches Justiz- und Polizeidepartement EJPD
Département fédéral de justice et police DFJP
Dipartimento federale di giustizia e polizia DFJP
Staatssekretariat für Migration SEM
Secrétariat d'Etat aux migrations SEM
Segreteria di Stato della migrazione SEM

Flüssigkeit und Korrektheit – Flüssigkeit oder Korrektheit?

Ergebnisse einer Interventionsstudie in niederschweligen Intensivkursen für erwachsene DaZ-Anfängerinnen und -Anfänger

Fluidité et correction – fluidité ou correction ?

Synthèse des résultats d'une étude expérimentale portant sur des cours intensifs à bas seuil d'allemand langue seconde, destinés à des adultes débutants

Fluena e correttezza – fluena o correttezza?

Risultati di uno studio d'intervento nei corsi di base intensivi di tedesco come lingua seconda per principianti

Fluency and Accuracy – Fluency or Accuracy?

Results of an intervention study in low-threshold intensive courses for adult beginners of German as a second language

Peter Lenz, Malgorzata Barras, Fabienne Manz

2017

Bericht des Wissenschaftlichen Kompetenzzentrums für Mehrsprachigkeit
Rapport du Centre scientifique de compétence sur le plurilinguisme
Rapporto del Centro scientifico di competenza per il plurilinguismo
Report of the Research Centre on Multilingualism

Deutsch Flüssigkeit und Korrektheit – Flüssigkeit oder Korrektheit? 5

	Zusammenfassung	6
1	Einleitung	8
2	Ausgangslage und theoretische Grundlagen	10
3	Ziele der Studie, Forschungsfragen	15
4	Methode	16
4.1	Untersuchungsdesign	16
4.2	Unterrichtsintervention	17
4.3	Datenerhebung	19
5	Ergebnisse	23
5.1	Testergebnisse	23
5.2	Ergebnisse aus den Interviews mit den Kursleitenden	29
6	Diskussion	31
7	Fazit	36
8	Bibliographie	144

Français Fluidité et correction – fluidité ou correction? 39

	Résumé	40
1	Introduction	42
2	Situation initiale et bases théoriques	44
3	Objectifs de l'étude et questions de recherche	50
4	Méthode	51
4.1	Conception de l'étude	51
4.2	Intervention didactique	52
4.3	Récolte des données	54
5	Résultats	58
5.1	Résultats des tests	58
5.2	Résultats des entretiens avec les enseignants	63
6	Discussion	66
7	Bilan	71
8	Bibliographie	144

Italiano Fluenza e correttezza – fluenza o correttezza? 75

	Riassunto	76
1	Introduzione	78
2	Situazione iniziale e basi teoriche	80
3	Obiettivi dello studio e domande di ricerca	85
4	Metodo	86
4.1	Design della ricerca	86
4.2	Intervento	87
4.3	Raccolta dei dati	89
5	Risultati	93
5.1	Risultati dei test	93
5.2	Risultati dei colloqui con gli insegnanti	99
6	Discussione	102
7	Conclusioni	107
8	Bibliografia	144

English Fluency and Accuracy – Fluency or Accuracy? 111

	Abstract	112
1	Introduction	114
2	Context	116
3	Study Aims, Research Questions	121
4	Methods	122
4.1	Study Design	122
4.2	Classroom Intervention	123
4.3	Data Collection	125
5	Results	129
5.1	Test Results	129
5.2	Results from Teacher Interviews	135
6	Discussion	137
7	Summary	141
8	Bibliography	144

Flüssigkeit und Korrektheit – Flüssigkeit oder Korrektheit?

Ergebnisse einer Interventionsstudie in niederschweligen Intensivkursen für erwachsene DaZ-Anfängerinnen und -Anfänger

—
Peter Lenz, Malgorzata Barras, Fabienne Manz¹

¹ Ein grosser Dank geht an Thomas Studer, den Verantwortlichen des Forschungsbereichs, für seine hilfreichen Vorschläge zur Durchführung des Projekts und zur Gestaltung dieses Berichts.

Zusammenfassung

Die vorliegende Studie geht der Frage nach, mit welchem Erfolg in Deutsch-als-Zweitsprache-Kursen für teilweise schulungsgewohnte erwachsene Deutsch-Anfänger/innen mit einem Unterrichtskonzept gelernt wird, das weitgehend auf dem Lexikalischen Ansatz (*Lexical Approach*) (Lewis, 1993; 1997), und damit dem Lernen von Chunks (Mehrwort-einheiten mit teilweise variablen Stellen), sowie auf regelmässigem Flüssigkeitstraining (Nation & Meara, 2002) beruht. In einem quasi-experimentellen Design mit Interventions- und Kontrollgruppe (IG und KG) arbeiteten zwölf Klassen in sechs Klassenpaaren jeweils einen ganzen Lehrwerkband durch. Unterricht und Materialien der einen Klasse jedes Paares setzten dabei das skizzierte Interventionskonzept um, Unterricht und Materialien der anderen Klasse das Kontrollkonzept, das im Wesentlichen mit dem Konzept der Originalmaterialien übereinstimmt. Wichtige Elemente des Konzepts der KG sind expliziter Grammatikunterricht und Wortschatzarbeit mit Einzelwörtern. Bei einem Eingangstest und einem Abschlusstest wurden verschiedene schriftliche und mündliche Sprachkompetenzen gemessen. Interviews mit den Kursleitenden sollten zudem u. a. Aufschluss geben über die Eignung der Unterrichtskonzepte für die Zielgruppe sowie über Besonderheiten der Unterrichts- und Lernprozesse in den beiden Gruppen.

Aus den Interviews geht hervor, dass nach beiden Konzepten insgesamt gut unterrichtet werden konnte. Kursteilnehmerinnen und -teilnehmer der IG mit Erfahrung im Fremdsprachenlernen sowie solche mit einem Fokus auf Korrektheit vermissten teilweise Grammatikunterricht und -materialien. Das Flüssigkeitstraining anerkannten sie insgesamt als gute Lerngelegenheit. Die Erfahrungen der Kursleitenden bestätigen, dass sich das Konzept der IG eher für Schulungsgewohnte und das Konzept der KG eher für Schulgewohnte eignet. Den Kursleitenden der IG fehlte gelegentlich bei der Fehlerkorrektur die grammatische Metasprache. Weiter hatten sie Mühe damit, dass beim Flüssigkeitstraining oft fehlerhafte Sprache wiederholt und somit eingeübt wurde, ohne dass sie genügend Gelegenheit für Korrekturen hatten.

Die Ergebnisse zu den verschiedenen sprachlichen Teilkompetenzen weisen zwischen Eingangstest und Abschlusstest fast alle einen Lernzuwachs aus. Ausnahmen sind einzig bei den Flüssigkeitsmassen zu finden. Bei der Kontrollgruppe zeigen die beiden wichtigen Flüssigkeitsindikatoren keinen Zuwachs an, bei der Interventionsgruppe dagegen schon. Die Fortschritte bei der *speech rate* (Anzahl Silben pro Sekunde der gesamten Sprechzeit unter Einschluss aller Pausen) und der *pruned speech rate* (entspricht der *speech rate*; Silben in gefüllten Pausen, Wiederholungen und Selbstkorrekturen werden aber nicht mitgezählt) geben einen starken Hinweis darauf, dass die Lernenden der IG die Aufgabe zur mündlichen Produktion mit weniger Planungs- und Überwachungsaufwand bewältigten als die Lernenden der KG. Die detaillierten Ergebnisse aus einem

weiteren mündlichen Testteil legen die Vermutung nahe, dass dem Gewinn an Flüssigkeit ein Verlust an Korrektheit gegenübersteht. In funktional-inhaltlicher Hinsicht unterscheiden sich die mündlichen Kompetenzen dagegen nicht. Im schriftlichen Teilttest finden sich ebenfalls keine Unterschiede zwischen den Gruppen, auch nicht im Bereich der Morphosyntax, wo eigentlich ein Unterschied zugunsten der KG erwartet wurde. Die differenzierten Befunde werden abschliessend vor dem Hintergrund von theoretischen Ansätzen zum Verständnis des Spracherwerbs und der Sprachproduktion diskutiert, und es werden Konsequenzen für die verglichenen Kurskonzepte und für die weitere Forschung vorgeschlagen.

1 Einleitung

Das Projekt „Der Einfluss von fremdsprachendidaktischen Verfahren auf das Sprachenlernen von erwachsenen Migrantinnen und Migranten in niederschweligen Sprachkursen“, das dieser Studie zugrunde liegt, wurde im Rahmen des Forschungsprogramms 2011-2015 des Kompetenzzentrums für Mehrsprachigkeit durchgeführt und vom Staatssekretariat für Migration (SEM) unterstützt. Es hatte zum Ziel, mit Hilfe von empirischen Befunden Licht in eine Frage zu bringen, die im Prozess der Erarbeitung des schweizerischen „Rahmencurriculums für die sprachliche Förderung von Migrantinnen und Migranten“ (Lenz, Andrey & Lindt-Bangerter, 2009) von Expertinnen und Experten diskutiert wurde. Es handelt sich um die Frage, ob im DaZ-Unterricht für oft schulungewohnte Kursteilnehmerinnen und -teilnehmer (KT) mit einem herkömmlichen didaktischen Ansatz gearbeitet werden soll, bei dem expliziter, analytischer Grammatikunterricht (*focus on forms* gemäss Long, 1991) und auf Einzelwörter fokussierende Wortschatzarbeit tragende Pfeiler sind, oder ob in dieser Art von Kursen eher unanalytisierte sprachliche Sequenzen (*Chunks*) gezielt gelernt und automatisiert werden sollen. Die vorliegende Studie versucht, evidenzbasierte Grundlagen zur Klärung dieser Frage bereitzustellen.

Zu diesem Zweck wurde ein quasi-experimentelles Design mit einer Interventions- und einer Kontrollgruppe umgesetzt. In sechs Klassenpaaren wurde die eine Klasse jeweils nach einem Chunk-orientierten Konzept und die andere nach einem stark Grammatik- und Einzelwort-orientierten Konzept unterrichtet, wie es vom Lehrwerk her eigentlich vorgesehen war.

Ursprünglich vier Anbieter von niederschweligen DaZ-Kursen arbeiteten mit uns zusammen. Bei drei Anbietern mit insgesamt zwölf Klassen konnte die Studie wie vorgesehen abgeschlossen werden. Der vierte Anbieter verliess nach einer ersten Phase das Projekt, weil die grosse Heterogenität in seinen Klassen eine Umverteilung der Kursteilnehmenden nötig machte. Die Zusammenarbeit verlief dank des Engagements aller Beteiligten sehr positiv.²

Einen wichtigen Pfeiler bei der Implementierung der Studie bildeten die Unterrichtsmaterialien. Für die Studie wurden die DaZ-Lehrwerke verwendet, welche die Anbieter auch sonst einsetzten. Zwei Anbieter verwendeten das gleiche Lehrwerk, der dritte ein anderes.³ Für die Umsetzung des Chunk-orientierten Konzepts mussten verschiedene Elemente der

2 Den folgenden Kursanbietern sei hiermit herzlich gedankt: aoz (Zürich), Formazione (Bern), Lernpunkt (Bern) und ECAP (Zentralschweiz/Luzern).

3 Es handelt sich um die für die Schweiz adaptierten DaZ-Lehrwerke *Schritte plus 1. Niveau A1/1. Ausgabe Schweiz* (Niebisch, Penning-Hiemstra, Specht & Bovermann, 2011) des Hueber Verlags und *Studio d A1. Schweizer Ausgabe* (Funk, Kuhn, Demme & Bayerlein, 2005) des Cornelsen Verlags.

beiden Lehrwerke gezielt abgeändert oder ergänzt werden. Schliesslich wurden die Materialien als gebundene Broschüren neu gedruckt.⁴

Untersuchungen wie die vorliegende sind selten. Zum einen ist das Sprachenlernen von erwachsenen Migrantinnen und Migranten noch wenig erforscht, obschon die Sprachförderung im Migrationsbereich zu einem auch gesellschaftlich relevanten Bereich geworden ist. Zum anderen sind quasi-experimentelle Interventionen, die mehrere Wochen dauern und Kurskonzeptionen insgesamt betreffen, selten. Sie stellen nicht nur in praktischer Hinsicht, sondern auch aus der Forschungsperspektive eine besondere Herausforderung dar – Letzteres deshalb, weil Faktoren, die den Einfluss von Interventionskonzepten potenziell verwässern, im Unterrichtsalltag schwerer zu kontrollieren sind als während punktueller Interventionen. Dafür erlauben Befunde aus Interventionen, die integraler Bestandteil des Ökosystems Unterricht sind, direktere Schlüsse im Hinblick auf eine eventuelle Umsetzung in der Kurspraxis.

4 Wir danken allen, die uns die Rechte zur Verwendung ihrer Materialien zur Verfügung gestellt haben.

2 Ausgangslage und theoretische Grundlagen

Im Zentrum der vorliegenden Studie steht ein Vergleich der Entwicklung der Sprachkompetenzen von erwachsenen Deutschanfängern, die in Intensivkursen für Deutsch als Zweitsprache (DaZ) nach zwei unterschiedlichen Konzepten unterrichtet wurden. Das allgemeine Ziel dieser Kurse besteht darin, die sprachliche Handlungsfähigkeit der KT zu erhöhen, damit sie besser am gesellschaftlichen Leben des Gastlandes teilnehmen können. Unsere Studie überprüft die kommunikative Handlungsfähigkeit allerdings nicht ganzheitlich durch kommunikative Testaufgaben, sondern fokussiert auf verschiedene linguistische Ressourcen und damit auf Teilaspekte kommunikativer Handlungsfähigkeit wie Sprechflüssigkeit und grammatische Korrektheit. Diese Schwerpunktsetzung bei den Ressourcen knüpft an Forschungstraditionen der Angewandten Linguistik an. Sie ermöglicht differenziertere und besser zu interpretierende Masse und rechtfertigt sich u.a. dadurch, dass bei Anfängern die Entwicklung eines linguistischen Repertoires als Grundlage für die Sprachverwendung im Vordergrund steht.

„Flüssigkeit“ in einer Sprache zu erlangen ist für viele oberstes Ziel des Sprachenlernens. Es bedeutet im Allgemeinen, dass jemand eine Sprache schnell, problemlos und korrekt verwenden kann (Segalowitz, 2005). Auf der kognitiven und physiologisch-artikulatorischen Ebene setzt Flüssigkeit Automatisierung voraus. Wenn ein komplexeres Verhalten neu erworben wird, erfordert es Aufmerksamkeit, Konzentration und Anstrengung. Die Performanz ist im Normalfall langsam und weist Fehler auf, sie verbessert sich aber durch Wiederholung, sodass schliesslich Handlungen leichter ausgeführt werden können (Aguado, 2002). Automatisierung hat nicht einfach mit der Beschleunigung von Sprachprozessen zu tun, sondern bedingt auch eine Restrukturierung der Vorgänge im Gedächtnis (Anderson u. a., 2004; Gatbonton & Segalowitz, 2005), die zu einer Entlastung des Arbeitsgedächtnisses bei der Sprachverwendung führt. Merkmale von Automatisierung sind höhere Geschwindigkeit, Mühelosigkeit, Unbewusstheit und Unaufmerksamkeit (Ballistizität) bei der Sprachverwendung, wobei diese Merkmale nicht immer gemeinsam auftreten müssen (Segalowitz & Hulstijn, 2009). Segalowitz & Hulstijn bezeichnen gerade die Ballistizität als kriteriales Merkmal für die Automatisierung von Sprache. Mit Ballistizität ist das Phänomen gemeint, dass automatisierte Sprachverwendung gegenüber Störungen relativ immun ist. Im Zusammenhang mit dem Erwerb einer L2 dürfte die Ballistizität von Automatismen dafür mitverantwortlich sein, dass sich bei Lernenden oft sprachliche Formen, die nicht zielsprachlichen Normen entsprechen, lange oder gar dauerhaft halten, d. h. fossilisieren⁵ (Bybee, 2008).

5 Long (2005) stellt die (endgültige) Fossilisierung als reales kognitives Phänomen in Frage. In seiner kritischen Analyse von Studien zeigt er, dass dieser gefürchtete defizitäre

Wenn Automatisierung bzw. Automtizität von Sprache für die sprachliche Entwicklung zentral ist, stellt sich für den Fremdsprachenunterricht die Frage, wie diese erreicht werden kann.

Aus der Sicht der *usage-based* („gebrauchsbasierten“) Linguistik geht es darum, beim Spracherwerb natürliche Prozesse in Gang zu bringen und zu nutzen, die letztlich zu automatisiertem Sprachgebrauch führen. Die Lehr- und Lernaufgabe besteht darin, in den Lernenden sog. *constructions* entstehen zu lassen und weiterzuentwickeln. Die Grundannahme besteht darin, dass Sprachenlernende (der L1 und weiterer Sprachen) Elemente und Konventionen des Sprachgebrauchs aus ihrer Kommunikationserfahrung ableiten. Die Erkenntnisse und Fertigkeiten aufgrund der Erfahrungen werden nach einem Abstraktionsprozess (Prozeduralisierung) im Gedächtnis als *constructions* gespeichert. Die *constructions* ordnen bestimmten (kommunikativen) Funktionen mental bestimmte sprachliche Formen zu (Ellis, 2005a). *Constructions* setzen sich aus festen (kognitiven) Chunks („*neuromotor routines*“) und einer variablen Anzahl austauschbarer Teile zusammen (Bybee, 2008). Bei *constructions* kann es sich um Wörter, Ausdrücke oder ganze Sätze handeln. Ein Beispiel für eine *construction* mit einer offenen Stelle für austauschbare Teile ist etwa der halb feststehende Ausdruck „Ich hätte gern ...“. *Constructions* müssen im Sprachmaterial erkannt und dann durch Wiederholung automatisiert werden, bevor sie in der Sprachverwendung zum Zuge kommen können. Sobald genügend automatisierte *constructions* zur Verfügung stehen, v. a. auch umfassendere *constructions*, die wiederum kleinere enthalten, kann flüssiger Sprachgebrauch entstehen. Die *constructions* entlasten dabei die limitierten Kapazitäten des Arbeitsgedächtnisses und geben kognitive Ressourcen beispielsweise für die Handlungssteuerung frei (Skehan, 2009).

In Bybee (2008) finden sich Überlegungen zu einem Curriculum, welches die Entwicklung von *constructions* fördert. Ein zentrales Anliegen besteht darin, die Erwerbsbedingungen so zu gestalten, dass der im Vergleich zur L1 geringere Kontakt mit der Zielsprache kompensiert wird. In einem ersten Schritt soll aufgrund von Korpusanalysen eine gezielte Auswahl von Sprachsequenzen getroffen werden. Mittels dieser Sequenzen sollen wichtige *constructions* einer Sprache anhand von prototypischen Sprachbeispielen vermittelt werden. Um Muster als solche erkennbar zu machen und den Erwerb

Endzustand des (fremdsprachlichen) Spracherwerbs oft wahrscheinlich da nicht vorliegt, wo er diagnostiziert wird. Nicht normgemässe sprachliche Merkmale weisen demnach primär auf einen aus Gründen unterschiedlicher Art nicht weitergeführten Spracherwerbsprozess hin.

in Gang zu bringen, sollen die Lernenden wiederholt mit den gleichen und gezielt mit abgeänderten Mustern konfrontiert werden.

Segalowitz & Hulstijn (2009) verweisen darauf, dass ein solcher oder ähnlicher Unterricht, der aus linguistischer Sicht für die Förderung der Automatisierung empfehlenswert wäre, in der Praxis auf Widerstand stossen dürfte: Er bietet zwar viel Wiederholung, um Automatisierung entstehen zu lassen, ist aber in Gefahr, daran zu scheitern, dass ihn die Betroffenen als dekontextualisiert und unattraktiv empfinden. Typischer kommunikativer Unterricht dagegen erscheint meistens inhaltlich relevant, umfasst aber zu wenig Wiederholung, um die Automatisierung genügend zu fördern. Eine Lösung sehen sie in Konzepten, die sich an den Kommunikationsbedürfnissen orientieren und vielseitig einsetzbare zielsprachliche Äusserungen vermitteln, die im Rahmen von Aktivitäten mit kommunikativen, inhaltlich bedeutsamen Aufgabenstellungen mit genügender Häufigkeit wiederholt werden.

Unsere Konzepte für Interventions- und Kontrollgruppe unterscheiden sich in didaktischer Hinsicht insbesondere durch einen anderen Umgang mit der Grammatik. In der Kontrollgruppe wird Grammatik explizit und systematisch unterrichtet und geübt (*focus on forms*), während sie in der Interventionsgruppe nicht als solche thematisiert wird. Damit kommt eine lange Zeit stark umstrittene Frage der Spracherwerbsforschung und der Fremdsprachendidaktik ins Spiel: Stehen Sprachkenntnisse, die im Unterricht explizit, aufgrund von Regeln, vermittelt und dann eingeübt werden, den Lernenden in gleicher Weise spontan zur Verfügung wie Kenntnisse, die sie selbst dem fremdsprachlichen Input entnommen haben? Bekannt ist v. a. die radikale Position von Krashen (1982), die besagt, dass dies nicht möglich sei (*non-interface position*) und regelbasierte Kenntnisse nur als sog. Monitor für die Überwachung der Sprachverwendung zur Verfügung ständen. Unterdessen hat sich die Auffassung durchgesetzt, dass komplexe Lernprozesse wie der Spracherwerb durchaus bei bewussten Erkenntnissen (deklarativem Wissen) ansetzen und schliesslich in automatisiertes Können übergehen können (nämlich als Elemente des prozeduralen Wissens; vgl. dazu das bekannte allgemeine Modell kognitiver Entwicklung von Anderson u. a., 2004). Eine grosse Metastudie zum Erfolg unterschiedlicher Unterrichtsformen (Norris & Ortega, 2001) gibt zudem starke Anhaltspunkte dafür, dass ein Unterricht, der systematisch oder punktuell auf sprachliche Phänomene fokussiert, erfolgreicher ist als bloss implizite Unterrichtsformen. Zu besonders grossen Fortschritten führt dabei offenbar ein Unterricht, der auch explizites Sprachwissen aufbaut. Ein neuerer Überblicksartikel von Spada & Tomita (2010), der mehr offene Testformen und mehr Ergebnisse zu implizitem Unterricht umfasst als die frühere Studie, bestätigt die Überlegenheit von Unterrichtstypen mit expliziten Anteilen gegenüber rein impliziten. Unser

Interventionskonzept umfasst zwar zahlreiche Übungsformen mit einem Fokus auf bestimmte sprachliche Phänomene, diese Phänomene werden aber nicht explizit thematisiert. Es bleibt also ein klarer Unterschied zum Konzept der Kontrollgruppe, das relativ viel explizite Grammatikarbeit vorsieht.

Für die Umsetzung des Interventionskonzepts stützen wir uns in erster Linie auf einen Ansatz aus dem Fachbereich der Fremdsprachendidaktik, nämlich den *Lexical Approach* von Michael Lewis,⁶ der seit den 1990-er Jahren Bekanntheit erlangt hat (z.B. Lewis, 1993; 1997; 2000). Wir haben ihn insbesondere durch Vorschläge von Nation (2001) zum Flüssigkeitstraining ergänzt.

Lewis selbst gründet seinen Ansatz v. a. auf Erkenntnisse der Korpuslinguistik zur Häufigkeit von mehr oder weniger festen Verbindungen zwischen Wortschatzelementen. Für Lewis sind lexikalische Elemente der Schlüssel zur Grammatik: „Language consists of grammaticalized lexis, not lexicalized grammar“ (Lewis, 1993: 34). Er empfiehlt, real vorkommende Mehrworteinheiten zum Ausgangspunkt des Lernens zu nehmen, weil Sprache zu einem grossen Teil aus solchen vorgefabrizierten Einheiten besteht, die nicht beliebig, d. h. gemäss dem grammatisch an sich Möglichen, manipuliert werden dürfen. Lewis nennt diese Einheiten Chunks. Dabei unterscheidet er zwei Hauptkategorien: Kollokationen (*collocations*) und Ausdrücke (*expressions*) (Lewis, 1997). Mit Kollokationen sind Wörter gemeint, die häufig in unmittelbarer Nachbarschaft auftreten (z. B. „Zähne putzen“ und nicht etwa „Zähne waschen“ wie in anderen Sprachen). Bei den Ausdrücken unterscheidet Lewis zwischen festen (*fixed*) Ausdrücken wie „Guten Morgen“ und halbfesten (*semi-fixed*) Ausdrücken, die dadurch charakterisiert sind, dass sie eine oder mehrere Stellen aufweisen, die Variation erlauben, z. B. „Ich hätte gern ...“.

In unserer Studie bezeichnen wir sowohl feste als auch halbfeste Ausdrücke gemäss Lewis als Chunks. Zu den festen Ausdrücken zählen wir auch funktionale Einwort-Einheiten wie „Danke!“. Kollokationen bilden im Interventionskonzept keine eigene Unterkategorie, sondern werden als Teile von Ausdrücken eingeführt.

Wer nach dem lexikalischen Ansatz unterrichtet, führt Chunks unterschiedlicher Länge bereits in frühen Lernstadien ein, allerdings ohne sie grammatisch zu analysieren (Lewis, 1993). Sie werden im Rahmen eines handlungs- und bedürfnisorientierten Unterrichts vermittelt, damit die Lernenden möglichst rasch über sprachliche Mittel verfügen, mit denen sie im Alltag kommunizieren können. Vom Chunk-basierten Unterricht wird erwartet, dass er die Sprachflüssigkeit fördert, weil er längere Sprachelemente zur

6 Weitere, etwas weniger bekannte Versuche, die Erkenntnisse zu den Chunks in der Sprache curricular umzusetzen, stammen von Willis (1990) und Nattinger & DeCarrico (1992).

Verfügung stellt, die als Ganzes abgerufen werden können. Dies ist ein Aspekt, der das Gelingen von Kommunikation begünstigt. Im Vergleich zum konventionellen Unterricht haben das Lernen und Üben von Satzgrammatik sowie das willkürliche Speichern von neuen Einzelwörtern wenig Bedeutung. Dafür werden Hör- und später Leseverstehen besonders trainiert sowie ein aktives Lernen mit einem personalisierten Chunk-Wörterbuch (*lexical notebook*) und weiteren Lernheften gepflegt (Lewis, 1997).

Der Lexikalische Ansatz wird durch die Fachliteratur verschiedentlich unterstützt. Der grosse Anteil von Chunks in der Sprache wird nicht bezweifelt. Wray & Perkins (2000) weisen darauf hin, dass die Sprache eines erwachsenen muttersprachlichen Sprechers (des Englischen) zu bis zu 70% aus formelhaften Ausdrücken (inkl. Kollokationen) besteht. Verschiedentlich sind auch Verweise auf den Nutzen des Lernens von Chunks für die kommunikative Sprachverwendung bereits in einer frühen Phase des Erwerbs zu finden. Werden Chunks wie Lexik gelernt, stehen sie bald als (korrekte) Kommunikationsmittel zur Verfügung, auch wenn ein Mangel an Wissen über das Funktionieren der darin enthaltenen Sprache besteht. Wenn die Kommunikation dank Chunks besser gelingt, kann dies zu mehr Chancen zum Weiterlernen führen (Aguado, 2002; Nation & Newton, 2009; Wray, 2002). Was die Entwicklung des impliziten und expliziten Sprach-Regelwissens aufgrund der erworbenen Chunks angeht, wird erwartet, dass die Lernenden nach und nach bzw. ab einem bestimmten Entwicklungszeitpunkt ihre mentale Sammlung formelhafter Ausdrücke auf ihre Regelhaftigkeit hin analysieren (z. B. Lewis, 1997; Wray, 2000). Die Ergebnisse der Analyse werden dann zusätzlich zum ursprünglichen Chunk gespeichert und für den Sprachgebrauch zur Verfügung gestellt (Wood, 2002).

Ob und wann diese Analyse stattfindet, ist eine entscheidende Frage. Bybee (2008) schlägt vor dem Hintergrund der Gebrauchsbasierten Linguistik vor, den Lexikalischen Ansatz durch Unterrichtselemente zu ergänzen, die gezielt morphosyntaktische *constructions* entwickeln helfen.

Im Unterrichtskonzept für die Interventionsgruppe wird der Lexikalische Ansatz durch gezieltes Flüssigkeitstraining ergänzt, um durch Wiederholung den Erwerb zu unterstützen. Diese Entscheidung gründet auch auf Paul Nations Konzept der „vier Standbeine“ (engl. *four strands*) eines guten Sprachunterrichts (z. B. Nation & Meara, 2002). Gemäss diesem Konzept sollen 25% der Unterrichtszeit für Sprachaktivitäten reserviert werden, in denen bereits bekannte Sprache ohne speziellen Sprachfokus, dafür unter leichtem Zeitdruck, verwendet werden. Dadurch soll sie anschliessend für den flüssigen Gebrauch zur Verfügung stehen. Zwei Varianten werden unterschieden: Bei der ersten wird bekannte Sprache mehrfach wiederholt (z. B. mit steigendem Zeitdruck), bei der zweiten ist die Wortschatzanreicherung das Ziel: Bekannte Lexik wird in neuen Kombinationen und Anwendungen erfahren (z. B. durch das Lesen langer, aber vereinfachter Texte). Im Interventionskonzept wurde nur der erste Aspekt umgesetzt.

3 Ziele der Studie, Forschungsfragen

Ein erstes Ziel der Studie bestand in der Entwicklung eines tragfähigen Kurskonzepts für die Interventionsgruppe sowie dessen Umsetzung durch die Adaptation von Unterrichtsmaterialien. Weiter mussten Sprachkursanbieter (KA) gefunden werden, die bereit waren, während zwei bis vier Monaten weitgehend nach unseren Vorgaben zu arbeiten. Das Hauptinteresse der Studie galt den Lernfortschritten der Kursteilnehmenden. Sie mussten mit geeigneten Instrumenten erfasst und anschliessend analysiert werden.

Die Studie sollte die folgenden Fragen beantworten:

- i. Gelingt es, gemäss den Grundideen des Interventionskonzepts ein die Kurse prägendes Unterrichtskonzept mit den entsprechenden Materialien zu schaffen und beides in niederschweligen DaZ-Kursen umzusetzen?
- ii. Stösst das Interventionskonzept bei Kursleitenden (KL) und Kursteilnehmenden (KT) auf Interesse und Akzeptanz?
- iii. Ist das Interventionskonzept v. a. für bildungsferne KT geeignet?
- iv. Welchen sprachlichen Lernerfolg in verschiedenen Teilbereichen verzeichnen die KT, die nach dem Interventionskonzept unterrichtet werden, im Vergleich zu den KT, die nach dem Kontrollkonzept unterrichtet werden?

Die Fragen i bis iii sollten aufgrund von Interviews mit den KL beantwortet werden.

Der Lernerfolg der KT wurde mittels mündlicher und schriftlicher Sprachtests erfasst. Im Hinblick auf den Vergleich der Ergebnisse der beiden Gruppen galt Folgendes (im Sinne von Hypothesen): Das Interventionskonzept würde dann als erfolgreicher als das Kontrollkonzept eingestuft, wenn die Interventionsgruppe (IG) im Vergleich zur Kontrollgruppe (KG) die folgenden Ergebnisse erreichen würde:

- iv.a. Häufigere Nutzung von zielsprachlichen Chunks
- iv.b. Flüssigerer Sprachgebrauch
- iv.c. Korrektheit im mündlichen Gebrauch eher höher (intuitive Korrektheit); im schriftlichen Sprachgebrauch gleich hoch oder nur geringfügig geringer (Korrektheit nach dem Einsatz von Regeln)
- iv.d. Vergleichbare Wortschatzkenntnisse (Bedeutungsebene) sowie vergleichbare Fähigkeit, Inhalte zu kommunizieren (ohne Berücksichtigung der formalen Korrektheit)

4 Methode

4.1. Untersuchungsdesign

Für die Studie wurde eine quasi-experimentelle Untersuchungsanlage mit einer Interventions- und einer Kontrollgruppe gewählt. Dabei wurde ausschliesslich mit niederschweligen⁷ Intensivkursen zusammengearbeitet, Kursen also, in welchen die KT während vier bis fünf Tagen, d. h. mindestens acht und maximal elf Stunden pro Woche Deutsch lernten. Für die Bearbeitung eines Lehrwerkbandes waren je nach Anbieter zwischen acht und 16 Unterrichtswochen vorgesehen.

Um Effekte durch solche und mögliche weitere anbieterspezifische Unterschiede auszugleichen, wurde bei jedem der Anbieter vor Kursbeginn aus den angemeldeten Deutschanfänger/innen⁸ jeweils ein Klassenpaar gebildet, das nach unseren beiden Konzepten unterrichtet wurde. Die Zuteilung zur Interventions- oder Kontrollgruppe erfolgte nach dem Prinzip einer geschichteten Zufallsstichprobe. Die spezifische Erfahrung der Kursanbieter mit der Zielgruppe wurde genutzt, um potenziell lernrelevante Merkmale in den beiden Klassen möglichst konstant zu halten. Dabei handelte es sich primär um die Merkmale Herkunft/Muttersprache und Schulgewohnheit.

Die so gebildeten Klassen wurden durch die Kursanbieter den Kursleitenden (KL) zugeordnet. Insgesamt konnten auf diese Weise bei vier verschiedenen Anbietern sieben Klassenpaare gebildet werden. Ein Klassenpaar musste wieder aufgelöst werden, weil die Zusammensetzung der Klassen wegen unserer Zuordnungsprinzipien zu heterogen war, was zu Unzufriedenheit führte und eine Einteilung in Niveaugruppen nötig machte.

Unsere Intervention erfolgte in zwei gleich langen Teilen, wobei die erste Hälfte den Charakter einer Vorphase hatte. In dieser Zeit erwarben viele KT ihre ersten Deutschkenntnisse, und die KL gewöhnten sich an die Arbeit mit unseren Konzepten und Materialien. Am Ende dieser ersten Phase wurde der Eingangstest für die Studie durchgeführt, der von der Struktur her mit dem Abschlusstest nahezu identisch war. Nach der ersten Kurshälfte kam es zu Abgängen und Neuzugängen bei den KT. Alle neu Dazugekommenen legten ebenfalls den Eingangstest ab. Normalerweise in der letzten Woche der zweiten Kurshälfte wurde dann der Abschlusstest durchgeführt. In der quantitativen Analyse spiegeln sich also die Lernergebnisse der zweiten Kurshälfte.

7 Bei niederschweligen Kursen ist der Zugang in verschiedener Hinsicht erleichtert, insbesondere hinsichtlich des Grades der schulischen und sprachlichen Vorkenntnisse und der Finanzierung.

8 Diese Bezeichnung bezieht sich auf die Einstufung der Kursinteressierten durch die Kursanbieter für einen Kurs auf Niveau A1.1 gemäss GER (Europarat & Goethe-Institut Inter Nationes etc., 2001).

4.2. Unterrichtsintervention

Die Intervention im Rahmen dieser Studie umfasste viele Aspekte der Kursgestaltung. Sie wurde in früheren Intensivkursen ausführlich erprobt und anschliessend verfeinert. In der Interventionsphase erfolgte die Umsetzung mittels Einführungs-Workshops mit KA und KL, Klassenbesuchen, Gesprächen, Unterrichtsleitfäden, Unterrichtsjournalen und vor allem auf der Basis der Unterrichtsmaterialien, die vieles vorgaben. Sowohl die Interventions- als auch die Kontrollgruppe waren davon betroffen, die Kontrollgruppe aber umfangmässig weniger.

Innerhalb eines Klassenpaares ist der Unterricht von IG und KG in verschiedener Hinsicht vergleichbar: Das Ausgangslehrwerk (und damit beispielsweise die Themen) ist identisch; es stehen dieselbe Stundenzahl und die gleiche Infrastruktur zur Verfügung, und die Dauer der Hausaufgaben wird zwischen den KL des jeweiligen Klassenpaares abgestimmt. Gemäss dem jeweiligen Konzept werden aber systematisch unterschiedliche Schwerpunkte gesetzt und teilweise andere Materialien verwendet.

Die Vorgaben durch das Konzept sind für die KL auch da verbindlich, wo sie eigene Aktivitäten einbringen. Grundsätzlich wurde versucht, von der Projektseite her genügend konzeptgerechtes Material mit Didaktisierungen zur Verfügung zu stellen, sodass Vermischungen nur im kleinen Rahmen vorkommen konnten.

Die beiden eingesetzten Lehrwerke wurden für die Interventionsgruppe aufwändig angepasst und neu gedruckt. Es wurden Elemente gestrichen, verändert oder ergänzt, damit die Arbeit mit Chunks zu einem dominierenden Element wurde. Die Änderungen an Konzept und Lehrwerk im Vergleich zur KG lassen sich folgendermassen zusammenfassen:

- Einführung eines täglichen Flüssigkeitstrainings von 20-30 Min. Dauer. Darin wird je zur Hälfte früher Gelerntes und am Vortag Gelerntes umgewälzt. Wichtig ist, dass die im Flüssigkeitstraining verwendeten sprachlichen Mittel bereits bekannt sind und zuvor im Unterricht geübt wurden. Oft werden beim Flüssigkeitstraining einfach die Aktivitäten aus dem Lehrwerk oder aus Zusatzmaterialien wiederholt, aber unter einem gewissen Zeitdruck⁹ oder zumindest im

9 Bereits die Erprobung hatte gezeigt, dass Zeitdruck bei gewissen Aktivitäten (wie z.B. dialogisches Sprechen) Stress erzeugt und negativ wahrgenommen wird. Deshalb wurden keine Aktivitäten eingefordert, bei denen der Zeitdruck im Vordergrund steht, wie etwa das von Nation empfohlene Wiederholen des gleichen mündlichen Beitrags in immer kürzerer Zeit.

Bewusstsein, dass Flüssigkeit im Vordergrund steht. Als vielfältig einsetzbare Übungsform wird das *Pointing* empfohlen: Ein Stichwort, eine Wortkarte oder ein Bild ist Anlass für Äusserungen unterschiedlicher Länge und Anzahl. Das Flüssigkeitstraining umfasst auch schriftliche Aktivitäten, z. B. das Schreiben von möglichst vielen Sätzen zu einem Bild als Gruppenwettbewerb.

- Ersetzen der Wortschatzlisten und auch der Grammatikübersichten durch ein *Lexical Notebook*, wie es von Lewis (1997) empfohlen wird. Dieses entsteht nach und nach aus dem Kurs heraus und enthält zahlreiche personalisierte Äusserungen für den eigenen kommunikativen Gebrauch.
- Ergänzung des Lehrwerks durch sog. Alltagsdialoge (aus Gründen der Stoffmenge nur beim einen Lehrwerk). Alltagsdialoge werden auch bei der KG ergänzt, aber anders didaktisiert. Vor allem werden die *turns* und Chunks aus den Dialogen in der IG intensiver behandelt. Die Alltagsdialoge werden in schematisierter Form und mit Lücken versehen auch als Faltblätter (Sprechkarten) zur Verfügung gestellt.
- Systematische Anpassung des Übungsgeschehens. Wo die KG mit Wörtern oder Morphemen arbeitet, arbeitet die IG mit Chunks. Übungstypen, die Lewis (1997) empfiehlt, bilden die Basis dafür:
 - Zuordnen – z. B. Bildung von kommunikativ relevanten Chunks aus Teilen
 - Sortieren – z. B. Ordnen „durcheinandergeratener“ Gesprächsbeiträge aus verschiedenen Gesprächen
 - Vervollständigen – z. B. Lückendiktate, wobei ganze Chunks ergänzt werden müssen oder relevante Chunks entstehen. Der Typ Vervollständigen wird im Zusammenhang mit den Alltagsdialogen mit der Zeit so weit getrieben, dass ganze Sprecherrollen ergänzt werden müssen.
- Ersatz der stark grammatiklastigen Lernzielkontrollen durch analoge Instrumente mit einem Chunk-Fokus.

Charakteristische Merkmale des Übungsgeschehens in den ursprünglichen Lehrwerken, d. h. auch im Unterricht der KG, wurden aus den Materialien der IG weitgehend eliminiert und ersetzt. Dabei handelt es sich insbesondere um Aktivitäten zum Erarbeiten von Grammatikregeln; Aktivitäten, bei denen grammatische Phänomene identifiziert und eingeordnet werden müssen; Einsetzübungen zur Morphologie und Syntax; grammatische Umformungen.

Die Materialien für die KG wurden nur geringfügig um einige Zusatzmaterialien ergänzt. Die KL wurden darauf verpflichtet, einen DaZ-Unterricht zu erteilen, der den „Geist“ der Lehrwerke, so wie er im Konzept für die KG interpretiert wird, möglichst genau umsetzt. Für sie hiess das, dass sie weniger eigene Materialien und Ideen einbringen konnten als sonst. Dafür war der Unterricht in den KG klarer definiert und über die verschiedenen Kursleitenden hinweg ähnlicher.

4.3. Datenerhebung

Drei Typen der Datenerhebung sind zu unterscheiden: 1) das Erfassen der Teilnehmerdaten aufgrund der Informationen der Kursanbieter, 2) die mündlichen und schriftlichen Tests zu zwei Zeitpunkten (Eingangs- und Abschlusstest zur zweiten Kurshälfte) sowie 3) die Interviews mit den Kursleitenden.

Teilnehmerdaten

Alle Kursanbieter erfassen die Personalien sowie eine Reihe von Hintergrunddaten, die aus administrativen Gründen oder für die Kurszuordnung von Interesse sind. Für die Gruppenbildung in unserer Studie waren v. a. der sprachlich-kulturelle und der Bildungshintergrund wichtig. Weitergehende, beispielsweise bereits auf Erfahrungen von früheren Kursleitenden beruhende Informationen zur wahrgenommenen Lerngewohntheit oder Lernfähigkeit waren nur punktuell vorhanden und konnten deshalb weder bei der Gruppenbildung noch bei der Datenanalyse (als Variablen, die unterschiedlichen Lernzuwachs erklären könnten) einbezogen werden.

Tests

Grundsätzlich wurde in allen Testteilen nur Wissen und Können getestet, das die Kursteilnehmenden in den Kursen erworben haben könnten. Mögliche Einflüsse aus dem privaten Umfeld sollten minimiert werden. Für jedes der beiden Lehrwerke wurde eine eigene Testversion erstellt. Der grösste Teil der Testitems war jedoch identisch. Der Abschlusstest bestand aus Items, die bereits im Eingangstest vorkamen, und aus solchen, die sich auf den zweiten Kursblock bezogen. Das Testdesign mit überlappenden Items war eine Bedingung dafür, dass alle Testergebnisse pro getestetem Kompetenzbereich auf jeweils einer einzigen Skala abgebildet werden konnten.

Geschulte Testleiter/innen führten die Tests vor Ort durch. Beim schriftlichen Test handelte es sich um eine Gruppenprüfung von jeweils rund 80 Minuten Dauer. Der mündliche Test wurde einzeln, betreut von einer Testleiterin oder einem Testleiter, vollständig am Computer absolviert. Der erste mündliche Test dauerte ca. 30-40 Min., der zweite ca. 35-45 Min. pro KT.

Die folgende Tabelle gibt einen Überblick über die getesteten Bereiche (Testkonstrukte) sowie die Testinstrumente, die zu den beiden Zeitpunkten eingesetzt wurden:

- 10 Der C-Test ist ein Lückentext-Format, bei dem ab dem zweiten Satz eines Textes bei jedem zweiten Wort die zweite Hälfte ergänzt werden muss (C-Test-Team, o. J.).
- 11 In unserer Umsetzung der Elicited-Imitation-Aufgaben (Erlam, 2009) am Computer (mit der Software PsychoPy, Peirce, 2009) hörten die KT jeweils ein Wort oder eine Wortsequenz, mussten dann eine Zahl vom Bildschirm lesen und sollten nach einem Signal das vorher Gehörte wiederholen. Dabei soll die Störung durch die Zahl verhindern, dass das Gehörte unverarbeitet aus dem sensorischen Gedächtnis (Shiffrin, 2003) sozusagen zurückkopiert werden kann.
- 12 Die *speech rate* und die beiden anderen Masse sind Indikatoren für die Flüssigkeit von mündlichen Äusserungen (*utterance fluency*) (Segalowitz, 2010). Indem die *speech rate* die reine Sprechgeschwindigkeit und die Pausenlänge zusammenfasst, ist sie ein technisches Flüssigkeitsmass, das gut mit der wahrgenommenen Flüssigkeit übereinstimmt. Für die reine Sprechgeschwindigkeit (*articulation rate*) gilt dies nicht (Kormos & Dénes, 2004; de Jong, Steinel, Florijn, Schoonen & Hulstijn, 2012).
- 13 Keine Artikulation während mehr als 250 Millisekunden wurde als stille Pause interpretiert.

Getesteter Bereich / Testkonstrukt	Testaufgabentyp [kodierter Aspekt]
Schriftlicher Test	
1a) Allgemeine schriftliche Sprachkompetenz (Inhaltsfokus)	C-Test ¹⁰ mit 4 Texten zu 15 Lücken [zumindest erkennbare Rekonstruktion des Originalwortes]
1b) Allgemeine schriftliche Sprachkompetenz (Inhalts- und Formfokus)	C-Test mit 4 Texten zu 15 Lücken [vollständige Rekonstruktion des Originalwortes]
2) Grammatik: Morphosyntax (potenziell regelgeleitet) Grammatik: Morphologie (potenziell regelgeleitet) Lexikogrammatik: Präpositionen (potenziell regelgeleitet)	Aus Elementen Sätze bilden Sätze variieren Plural angeben Sätze ergänzen
Mündlicher Test	
3a) Mündliches Sprachverstehen und mündliche Sprachproduktion (Inhaltsfokus)	<i>Elicited Imitation</i> ¹¹ [Anzahl wiedergegebener Inhaltselemente] <i>Elicited Imitation</i> [Anzahl wiedergegebener Wortkerne]
3b) Mündliches Sprachverstehen und mündliche Sprachproduktion (Inhalts- und Formfokus)	<i>Elicited Imitation</i> [Anzahl vollständig wiedergegebener Wörter] <i>Elicited Imitation</i> [Anzahl vollständig oder im Kern wiedergegebener <i>Chunks</i>]
4) Formelhaftigkeit mündlicher Äusserungen (<i>Chunking</i>)	<i>Elicited Imitation</i> [Vollständigkeit und Genauigkeit der <i>Chunks</i>]
5) Grammatik: Morphologie (vorwiegend intuitionsgeleitet)	<i>Elicited Imitation</i> [Korrektheit der Morphosyntax]
6) Sprechflüssigkeit (kognitive Flüssigkeit)	Gesteuerte mündliche Produktion; Steuerung durch Bilder und Tonaufnahmen als Sprachvorbilder [a. <i>speech rate</i> ¹² – Anzahl Silben pro Sekunde in einem festen Ausschnitt der Aufnahme (unter Einschluss aller Pausen) ¹³] [b. <i>pruned speech rate</i> – Anzahl Silben pro Sekunde in einem festen Ausschnitt der Aufnahme (unter Einschluss aller Pausen), ausschliesslich der Silben in gefüllten Pausen, Wiederholungen und Selbstkorrekturen] [c. <i>articulation rate</i> – Anzahl Silben pro Sekunde allein in den Teilen des Ausschnitts, in denen tatsächlich gesprochen wird (keine stillen Pausen)]

Tabelle 1: Überblick über Testkonstrukte und Testinstrumente beim Eingangs- und Abschlusstest

Interviews mit den Kursleitenden

Die Zusammenarbeit mit den Kursleitenden wurde nach Abschluss der Experimentalphase mit einem Interview abgeschlossen. Die Interviews für Unterrichtende der Interventions- und der Kontrollgruppe unterschieden sich nur hinsichtlich konzeptspezifischer Fragen. Zwei übergeordnete Themenbereiche wurden behandelt: 1) das Konzept und 2) die Teilnehmenden. Die Interviews dauerten zwischen 30 und 100 Min.

Die 16 Interviews wurden transkribiert und mittels einer Inhaltsanalyse (Mayring, 2008) nach Gesichtspunkten ausgewertet, die im Hinblick auf die Fragestellungen der Studie relevant sind.

- i. Umsetzung des Konzepts
 - a. Vergleich zum üblichen Unterricht der KL
 - b. Konkrete Umsetzung des Konzepts
 - c. Einflussfaktoren
- ii. Die Arbeit mit den KT bzw. der Klasse
- iii. Vor- und Nachteile des Konzepts
 - a. Sicht der KL
 - b. Sicht der KT (nach Einschätzung der KL)

Aufbereitung der Testleistungen

Die schriftlichen und mündlichen Leistungen der Kursteilnehmenden wurden jeweils von einem Kodierer-Tandem auf der Basis von ausführlichen schriftlichen Richtlinien für die Datenanalyse aufbereitet.

Aufwändig waren die Arbeiten zur Ermittlung der Flüssigkeitsindizes aufgrund von monologischen Sprechsequenzen. Die Tonaufnahmen wurden mithilfe der Software *Praat* (Boersma & Weenink, o. J.) und ergänzend von Hand segmentiert, kategorisiert und ausgezählt. Daraus konnten dann die Indizes berechnet werden.

5 Ergebnisse

5.1. Testergebnisse

Die Ergebnisse der Kursteilnehmenden in den schriftlichen Testteilen und beim *Elicited-Imitation*-Test wurden Rasch-skaliert.¹⁴ Dadurch konnten die Ergebnisse aufgrund der vier eingesetzten Testvarianten (zwei Lehrwerke, zwei Zeitpunkte) auf dieselbe Skala gebracht und so vergleichbar gemacht werden.¹⁵ Die Indikatoren zur mündlichen Flüssigkeit konnten direkt verwendet werden, weil sie sich beim Eingangs- und beim Ausgangstest auf dieselbe Aufgabe zur mündlichen Produktion beziehen.

Die statistische Auswertung basiert auf einfaktoriellen (Zugehörigkeit zur IG oder zur KG) Kovarianzanalysen (ANCOVA)¹⁶ mit Messwiederholung (Tests zu den beiden Zeitpunkten t1 und t2). Die Umsetzung erfolgte in Form von gemischten Modellen (mixed models).

Bei der Auswertung der verschiedenen Testteile musste jeweils eine leicht unterschiedliche Anzahl KT (N) einbezogen werden. Von 153 KT, die den zweiten Kursblock in Angriff nahmen, lagen schliesslich von 141 KT Daten aus mindestens einem Testteil vor, bei 129 Kandidaten aus allen Testteilen. Zwischen den Testergebnissen beim Eingangstest und dem Aufgeben des Kurses besteht kein signifikanter Zusammenhang.¹⁷ Tabelle 2 vermittelt einen Überblick über den Datenbestand. Das Fehlen von Testdaten ist v. a. auf organisatorisch bedingte Abwesenheiten der Probanden bei Testteilen zurückzuführen.

14 Die statistischen Analysen wurden mit Paketen der Software-Umgebung R ausgeführt (Chambers, 2008; Bates, Mächler, Bolker & Walker, 2015; Kiefer, Robitzsch & Wu, 2015; Lenth, 2015; Halekoh & Højsgaard, 2014).

15 Die Ergebnisse der Personen wurden als sog. WLEs (Warm, 1989) ausgegeben. WLE-Skalen überschätzen die tatsächlich vorhandene Variabilität in den Daten etwas. Der Einfachheit halber wurden die WLEs aber wie beobachtete Werte behandelt.

16 Anders als in der Varianzanalyse werden dabei auch Variablen (Kovariaten) einbezogen, von denen bekannt ist, dass sie individuelle Unterschiede beeinflussen können. In unserem Fall handelte es sich um Herkunft, Aufenthaltsdauer in der Schweiz, Schulbildung und Geschlecht. Hinzu kommt die Variable „Klassenpaar“, die dem Umstand Rechnung trägt, dass die Kursanbieter jeweils gleichzeitig aus einer grösseren Gruppe von Kursinteressierten eine Interventions- und eine Kontrollklasse bildeten, die dann unter sehr ähnlichen äusseren Bedingungen lernten.

17 Dies wurde mittels Pearson-Korrelationen zwischen der Variablen für „Kursaufgabe – ja oder nein“ und den Testergebnissen zu t1 überprüft.

	IG	KG	Total
Kursblock 2 begonnen	73	80	153
Kursblock 2 beendet (inkl. mind. 1 Testteil)	66	75	141
Frauen/Männer am Ende von Kursblock 2	34/32	34/41	68/73
Anzahl Schuljahre: Mittelwert (Min.-Max.)	8.8 (0-17)	9.3 (0-16)	9.1 (0-17)
Alter: Mittelwert (Min.-Max.)	33.2 (16-51)	32.3 (17-57)	32.7 (16-57)
Alle Testteile und Teiltests zu t2 abgelegt; Daten nutzbar	63	66	129
Schriftlichen Testteil zu t2 abgelegt	66	73	139
Mündlichen Teilttest I (<i>Elicited Imitation</i>) zu t2 abgelegt	63	68	131
Mündlichen Teilttest II (Mdl. Produktion/Flüssigkeit) zu t2 abgelegt	64	71	135

Tabelle 2: Übersicht über Kursteilnehmende und Testdaten

Tabelle 3 stellt die Testergebnisse im Überblick dar. Zu den drei Testteilen werden die Ergebnisse zu insgesamt elf untersuchten Bereichen dargestellt. Dabei werden jeweils die Mittelwerte pro Gruppe und Zeitpunkt angegeben. In der Klammer steht der Standardfehler. Verschiedene Signifikanztests überprüfen die Verallgemeinerbarkeit der Ergebnisse aus statistischer Sicht. Getestet werden für jeden untersuchten Bereich a) der Zuwachs von t1 nach t2 in IG und KG und b) der Unterschied des Zuwachses in IG und KG. Die Angaben zu diesem Zuwachsvergleich in der letzten, mittelgrau hinterlegten Spalte sind für die vorliegende Untersuchung zentral.¹⁸

18 Für die mündlichen Testteile sind jeweils zwei unterschiedliche p-Werte angegeben, einer („p“), der den Umstand ausser Acht lässt, dass dieselben Leistungen mehrfach, auf unterschiedliche Phänomene hin, kodiert und ausgewertet wurden, und ein zweiter („pH“, „H“ für Holm-korrigiert), welcher der Einhaltung des üblichen globalen 95%-Konfidenzniveaus pro Testteil Rechnung trägt. Die Holm-korrigierten p-Werte sind demnach seltener signifikant.

Untersuchter Bereich	Interventionsgruppe (IG)			Kontrollgruppe (KG)			IG vs. KG
	Eingangstest (t1)	Abschluss-test (t2)	Signifikanz des Zuwachses ¹⁹	Eingangstest (t1)	Abschluss-test (t2)	Signifikanz des Zuwachses	Unterschied im Zuwachs; Signifikanz ²⁰
Schriftlicher Testteil N = 139, n (IG) = 66, n (KG) = 73							
1a) Allgemeine schriftliche Sprachkompetenz: schriftlicher C-Test / vollständige Rekonstruktion	0.064 (0.126)	0.434 (0.126)	t = 5.13 p < 0.0001	0.134 (0.113)	0.540 (0.113)	t = 5.92 p < 0.0001	F = 0.13 ²¹ p = 0.719
1b) Allgemeine schriftliche Sprachkompetenz: schriftlicher C-Test / richtiges Wort	0.105 (0.081)	0.497 (0.081)	t = 4.22 p = 0.0003	0.077 (0.074)	0.448 (0.074)	t = 4.19 p = 0.0003	F = 0.03 p = 0.864
2) Gramm. Korrektheit im Schriftlichen (potenziell regelgeleitet): Anwendungsübungen zu Morphologie und Syntax	0.017 (0.181)	0.469 (0.181)	t = 3.79 p = 0.0013	0.386 (0.163)	0.829 (0.163)	t = 3.91 p = 0.0008	F = 0.003 p = 0.953
Mündlicher Testteil I: Elicited Imitation N = 131, n (IG) = 63, n (KG) = 68							
3a) Verständnis und Wiedergabe gesprochener Sprache: Anzahl wiedergegebener Inhaltselemente	-0.143 (0.131)	0.236 (0.131)	t = 5.74 p < 0.0001	-0.159 (0.119)	0.301 (0.119)	t = 7.24 p < 0.0001	F = 0.776 p = 0.380 p _H = nsig.
3b) Verständnis und Wiedergabe gesprochener Sprache: Anzahl vollständig wiedergegebener Wörter	0.055 (0.110)	0.379 (0.110)	t = 5.89 p < 0.0001	0.025 (0.099)	0.495 (0.099)	t = 8.88 p < 0.0001	F = 3.64 p = 0.059 p _H = nsig.
3c) Verständnis und Wiedergabe gesprochener Sprache: Anzahl wiedergegebener Wortkerne	-0.126 (0.135)	0.229 (0.135)	t = 5.42 p < 0.0001	-0.164 (0.122)	0.276 (0.122)	t = 6.99 p < 0.0001	F = 0.89 p = 0.348 p _H = nsig.
4) Formelhaftigkeit mündlicher Äußerungen: <i>Chunking</i>	-0.099 (0.119)	0.098 (0.119)	t = 3.27 p = 0.007	-0.160 (0.108)	0.205 (0.108)	t = 6.31 p < 0.0001	F = 4.075 p = 0.046 p _H = nsig.
5) Gramm. Korrektheit (eher intuitionsgeleitet): Korrektheit der Morphosyntax	-0.029 (0.142)	0.248 (0.142)	t = 3.50 p = 0.0035	-0.196 (0.129)	0.455 (0.129)	t = 8.545 p < 0.0001	F = 11.566 p = 0.0009 p _H = 0.0045
Mündlicher Testteil II: Mündl. Produktion N = 135, n (IG) = 64, n (KG) = 71							
6a) Sprechflüssigkeit: Anzahl Silben pro Aufnahmezeit in sec. <i>speech rate</i>	1.490 (0.054)	1.610 (0.054)	t = 2.73 p = 0.035	1.527 (0.049)	1.516 (0.049)	t = -0.29 p = 0.99	F = 4.765 p = 0.031 p _H = 0.093
6b) Sprechflüssigkeit: Anzahl Silben pro Aufnahmezeit in sec.; ohne Silben in Reparaturen, Wiederholungen und gefüllten Pausen; <i>pruned speech rate</i>	1.190 (0.053)	1.331 (0.053)	t = 3.27 p = 0.0075	1.214 (0.048)	1.230 (0.048)	t = 0.39 p = 0.98	F = 4.42 p = 0.037 p _H = nsig.
6c) Sprechflüssigkeit: Anzahl Silben pro Aufnahmezeit in sec. unter Ausschluss der stillen Pausen <i>articulation rate</i>	3.006 (0.057)	3.075 (0.057)	t = 1.35 p = 0.53	3.118 (0.051)	3.086 (0.051)	t = -0.65 p = 0.92	F = 2.02 p = 0.157 p _H = nsig.

Tabelle 3: Überblick über die Testergebnisse

Unter Berücksichtigung der Holm-Korrektur („p_H“-Werte) besteht nur bei einer Variablen ein Unterschied zwischen dem Zuwachs in der IG im Vergleich zur KG, nämlich im Bereich 5, bei der Korrektheit der Morphosyntax im mündlichen Testteil (*Elicited Imitation*). Dort ist der Zuwachs in der KG im Mittel grösser als in der IG. Ohne Berücksichtigung des mehrfachen Testens („p“-Wert) sind die folgenden Zuwachsvergleiche signifikant bzw. grenzsignifikant (Kursivschrift):

- zugunsten der KG: 3b) Anzahl vollständig wiedergegebener Wörter sowie 4) *Chunking*
- zugunsten der IG: 6a) *speech rate* sowie 6b) *pruned speech rate*

In drei von fünf ausgewerteten Bereichen der Leistungen in der *Elicited Imitation* zeigen sich also Vorteile für die Kontrollgruppe, in den beiden anderen ist der Zuwachs in den beiden Gruppen vergleichbar. Interessant ist ein Vergleich aus inhaltlicher Perspektive: Die Zunahme bei der KG ist dort stärker, wo Korrektheit und Vollständigkeit im Vordergrund stehen (Bereiche 3b, 4 und v. a. 5). Wo formalsprachliche Aspekte in den Hintergrund treten (3a, 3c) sind die Unterschiede zwischen IG und KG vernachlässigbar.

Der Zuwachs vom Eingangstest zum Abschlusstest ist *innerhalb* der beiden Gruppen fast immer signifikant. Ausnahmen (vier hellgrau schattierte Zellen) sind bei den Flüssigkeitsmassen zu finden: Bei der KG hat sich zwischen t1 und t2 keines der drei Flüssigkeitsmassen verändert, bei der IG nur die *articulation rate* nicht, das Mass für die reine Sprechgeschwindigkeit unter Ausschluss der stillen Pausen. Die Mittelwerte für die *speech rate* und die *pruned speech rate* haben sich dagegen signifikant erhöht, und zwar um 8 bzw. 12%.

19 p-Werte zum t-Test beim Zuwachs nach Gruppe mit der Korrektur nach Tukey für mehrfaches Testen.

20 p-Wert zum F-Test mit Korrektur für mixed-models (Kenward-Roger-Approximation).

21 Zugunsten der Übersichtlichkeit wird auf die Angabe von Freiheitsgraden verzichtet. Sie können im Wesentlichen aus den Gruppengrössen abgeleitet werden.

5.2. Ergebnisse aus den Interviews mit den Kursleitenden

Im Folgenden werden wichtige Ergebnisse aus den Abschlussinterviews mit den Kursleitenden berichtet. Die Darstellung folgt den in Abschnitt 4.3 genannten Gesichtspunkten.

Umsetzung des Konzepts durch die Kursleitenden

Um die Vergleichbarkeit des Unterrichts in allen Klassen der IG bzw. KG zu gewährleisten, wurden für beide Gruppen Konzepte erarbeitet, eingeführt und in Form von Leitfäden festgehalten. Die Umsetzung wurde zudem durch Beratung, Unterrichtsbesuche und Einsicht in die Kursleiterjournale begleitet. Aus den Interviews lässt sich schliessen, dass der Unterricht denn auch weitgehend durch das jeweilige Konzept geprägt war. Eigene Materialien und konzeptfremde Unterrichts- oder Lernaktivitäten spielten offenbar eine weit untergeordnete Rolle. Die in der IG als innovatives Element verlangten 20 bis 30 Min. Flüssigkeitstraining zu Beginn jedes Unterrichtstages wurden weitgehend umgesetzt, was aber v. a. zu Beginn oft nicht leicht fiel.

Aufnahme der Konzepte durch die Kursteilnehmenden und Klassen

Das Interventionskonzept wurde in den Klassen unterschiedlich aufgenommen. Die dynamischeren Sozialformen bei einigen mündlichen (Wiederholungs-)Übungen stiessen bei bestimmten Teilnehmenden (z. B. älteren) auf Widerstand, bei anderen waren sie beliebt. Eine Kursleiterin sieht in der vermehrten Interaktion den Grund für das besonders konstruktive Klima in ihrer Klasse.

Das Konzept der IG wurde offenbar von schulungewohnten und schwächeren Lernenden besser aufgenommen als von schulgewohnten KT oder solchen, die ihre Korrektheit verbessern wollten. Für Schulungewohnte und Schwächere war der Wegfall des Grammatikunterrichts eine Erleichterung. Schulungewohnte bezweifelten teilweise den Nutzen eines Flüssigkeitstrainings mit keiner oder wenig Fehlerkorrektur. Auch verlangten sie nach formellem Grammatikunterricht oder entsprechenden Materialien. Die regelmässige Wiederholung von Bekanntem im Rahmen des Flüssigkeitstrainings wurde aber auch als Lernchance erkannt.

In den Klassen der KG gab es vor allem wegen des von den Lehrwerken so vorgesehenen Wortschatzlernens auf der Grundlage von Wortschatzlisten negative Reaktionen.

Auf viel positives Echo in beiden Gruppen stiessen die speziell für die Zwecke dieser Studie angefertigten Alltagsdialoge auf CD, die in zwei konzeptspezifischen Didaktisierungsvarianten als Zusatzmaterial abgegeben wurden.

Einschätzung der Vor- und Nachteile der Konzepte durch die Kursleitenden

Mehrere Kursleitende beider Gruppen fühlten sich durch die Verpflichtung auf feste Konzepte und vorgegebene Materialien in ihren Möglichkeiten grundsätzlich eingeschränkt. Vor allem die Kursleitenden, die mit dem weniger spezifisch auf die Zielgruppe ausgerichteten Lehrwerk arbeiteten, hätten in Bezug auf Themen, Wortschatz und Grammatik gerne mehr didaktische Elemente weggelassen und diese durch eigene ersetzt. Umgekehrt waren die Kursleitenden der IG klar der Ansicht, dass die Umsetzung des Interventionskonzepts im Unterricht ohne angepasste Unterrichtsmaterialien nicht möglich gewesen wäre.

Die Kursleitenden der IG äusserten sich in den Interviews zu den verschiedenen Komponenten des Interventionskonzepts teils bestätigend, teils kritisch. Der grössere Stellenwert der Mündlichkeit, bedingt v. a. durch das Flüssigkeitstraining, wird begrüsst. Nicht sicher sind sich die Kursleitenden, ob die Förderung der Schriftlichkeit zu kurz kommt. Als Problem wurde die Individualisierung während des Flüssigkeitstrainings genannt: Die Kursteilnehmenden würden oft fehlerhaft sprechen und diese Fehler während längerer Zeit einüben, ohne dass die Kursleitenden wesentlich eingreifen konnten.

Die Übungsaktivitäten mit Chunks, die in der IG das grammatikzentrierte Übungsgeschehen ersetzen, werden insgesamt positiv beurteilt. Gleichzeitig wird das Fehlen von explizitem Grammatikunterricht problematisiert: Schulgewohnte haben entgegen ihrer Erwartung keinen Grammatikunterricht und keine entsprechende Materialgrundlage. Den Kursleitenden fehlt zudem im Unterricht die grammatische Metasprache als effizientes Werkzeug für Fehlerkorrektur und Erklärungen.

Das individualisierte *Lexical Notebook* (LN) wird grundsätzlich positiv beurteilt. Eingeschränkend wird aber auf den Korrekturaufwand verwiesen, der sich daraus ergibt, dass die Lernenden selbst Einträge machen, die wiederum korrigiert werden müssen, weil das LN als Grundlage für das individuelle Wiederholen vorgesehen ist. Schulungsgewohnte benötigten zudem (zu) viel Anleitung für die Arbeit mit dem LN. Positiv wird vermerkt, dass individualisierende Elemente wie das LN oder die Sprechkarten mit den Alltagsdialogen die Teilnehmenden zu mehr Lernerautonomie hinführen.

Anders als das Konzept der Interventionsgruppe wird das Konzept der Kontrollgruppe (d.h. der Unterricht nach den Vorschlägen aus dem Lehrerhandbuch) von den Kursleitenden als Konzept für Schulungsgewohnte und Schnelle bezeichnet. Die Bewertung dieses Umstands variiert jedoch. Es wird von zu viel Grammatik und zu vielen Grammatikbegriffen gesprochen, auch von zu viel expliziter Wortschatzarbeit. Es wird aber auch gesagt, dass diese intensive Wortschatzarbeit wichtig sei. Eine Kursleitende argumentiert, dass das Konzept der KG beibehalten werden sollte, um auch Schulungsgewohnte schulungsgewohnt zu machen.

6 Diskussion

Die Ergebnisse der Untersuchung ergeben insgesamt ein reichhaltiges Bild, sowohl zum Einsatz der beiden Konzepte im Unterricht als auch zur sprachlichen Entwicklung der Kursteilnehmenden.

In den Leitfadeninterviews findet sich eine ganze Reihe von Anhaltspunkten zu den ersten drei Forschungsfragen (vgl. Kap. 3). Das neue Unterrichtskonzept und die entsprechenden Materialien (Forschungsfrage 1) wurden offenbar erfolgreich entwickelt und eingeführt. Einzelne Kursleitende beider Gruppen bedauerten in den Interviews die starke Steuerung des Unterrichts während der Interventionsphase und hätten gerne mehr eigene, direkt an den Bedürfnissen der KT orientierte Materialien eingesetzt. Aus der Forschungsperspektive war eine solche Steuerung aber notwendig, und zwar bei beiden Gruppen. Anders wären die Ergebnisse weniger sicher auf den Einfluss der beiden Konzepte zurückzuführen.

Die Akzeptanz des Interventionskonzepts bei den Betroffenen (Forschungsfrage 2) ist gemischt und muss differenziert beurteilt werden. Da es bestimmte Elemente des sonst in unserem Kontext üblichen Fremdsprachenunterrichts ausschliesst (v. a. die explizite Beschäftigung mit Grammatik) und dafür andere stark betont (v. a. das tägliche Flüssigkeitstraining und ein Übungsgeschehen, bei dem Chunks in funktional sinnvollen Zusammenhängen geübt, aber nicht explizit analysiert werden), entspricht es nicht den Erwartungen von Lernenden mit Erfahrung mit herkömmlichem Fremdsprachenunterricht und von Lernenden, die den Kurs mit dem Ziel besuchen, die Korrektheit ihrer Sprache zu verbessern. Das Interventionskonzept fordert zudem die bisherige Praxis der Kursleitenden heraus. Ein Problem für KL und einen Teil der KT bestand offenbar darin, dass beim Flüssigkeitstraining Sprache automatisiert wurde, die sie noch nicht genügend beherrschten. Die Idee beim Flüssigkeitstraining ist eigentlich, dass nur Sprache verwendet wird, die bekannt ist, d. h. im Wesentlichen beherrscht wird (Nation & Meara, 2002). Offenbar stand aber den KT für die verlangten Leistungen zu wenig Sprache zur Verfügung, die sie sicher beherrschten, sodass zahlreiche Fehler gemacht wurden. Als Problem wird auch das Fehlen von explizitem Grammatikunterricht angesprochen. Schulungsgewohnte KT verlangten zum Teil nach Grammatikunterricht und/oder -materialien, während Kursleitende bei Feedbacks und Korrekturen die grammatische Terminologie vermissten. Schulungsgewohnten und schwächeren Lernenden dagegen kam der Verzicht auf explizite Grammatik offenbar entgegen und verkleinerte die Lernlast. Überhaupt kann die 3. Forschungsfrage zur Eignung des Interventionskonzepts für schulungsgewohnte KT aufgrund der Einschätzungen der KL insgesamt positiv beantwortet werden. Eine der didaktischen Herausforderungen von niederschweligen Kursangeboten besteht generell in der Heterogenität der KT. Dies manifestierte sich in unserer Studie besonders

an den Haltungen zum fehlenden expliziten Grammatikunterricht. Als mögliche Konsequenz könnte das Interventionskonzept spezifischer bei Schulungsgewohnten eingesetzt werden, oder es könnte durch einzelne Elemente aus dem konventionellen Konzept ergänzt werden, insbesondere Grammatikarbeit, die u. a. eine Terminologie für Korrekturhinweise zur Verfügung stellt und individuelles Weiterarbeiten mit den Materialien erlaubt. Umgekehrt könnte das spezifische Flüssigkeitstraining im Sinne von Nation & Meara (2002) in das konventionelle Konzept integriert werden, besonders weil bei einem stärkeren Korrektheitsfokus die Gefahr des Einübens von Fehlern kleiner würde und auch weil die gezielte Entwicklung von Sprachflüssigkeit in konventionellen Konzepten kaum vorgesehen ist, obschon Flüssigkeit als wichtiger Entwicklungsbereich und zentraler Kompetenzaspekt anerkannt ist.

Die Ergebnisse zu den Sprachtests entsprechen teilweise den (optimistischen) Erwartungen, die wir aufgrund von Aussagen in der Literatur zu den Konsequenzen eines Chunk-basierten Ansatzes formulierten (s. o. Abschnitt 3).

In drei Bereichen wurden die Erwartungen bestätigt:

- Der Zuwachs bei der Flüssigkeit ist in der IG grösser.
- Im Zuwachs bei den Wortschatzkenntnissen (Bedeutungsebene) sowie in der Fähigkeit, Inhalte zu kommunizieren sind wie erwartet keine Unterschiede zu erkennen.
- Der Zuwachs bei der schriftlichen Korrektheit unterscheidet sich nicht.

In zwei weiteren Bereichen wurden die Erwartungen nicht bestätigt:

- Bei der Verwendung von möglichst korrekten und vollständigen Chunks ist der Fortschritt in der IG kleiner als in der KG.
- Der Zuwachs bei der grammatischen Korrektheit im mündlichen Gebrauch ist in der KG grösser.

Interessant ist, die Ergebnisse gesamthaft und vor dem Hintergrund der Ergebnisse der Interviews mit den KL zu interpretieren.

Es fragt sich, wie weit das Gesamtbild der Testergebnisse mit der Beobachtung der KL zu tun hat, dass beim Flüssigkeitstraining (d. h. bei häufig vorkommenden Aktivitäten mit ausgeprägtem Wiederholungscharakter) oft fehlerhafte Sprache umgesetzt wurde: Werden durch die Wiederholung von nicht-korrekt-er Sprache nicht-zielsprach-

liche Chunks oder *constructions* gespeichert und abgerufen? Die Testergebnisse widersprechen dem zumindest nicht: Die mündlichen Äusserungen der IG sind im Mittel fehlerhafter als diejenigen der KG. Darauf verweisen die Vergleiche anhand der *Elicited Imitation* zu den drei Bereichen mit einem Korrektheitsanteil: grammatische Korrektheit, Präzision beim Wortschatz und korrektes Chunking. Gleichzeitig hat sich die mündliche Flüssigkeit in der IG im Unterschied zur KG verbessert, und zwar nicht durch eine Verbesserung der reinen Artikulationsgeschwindigkeit (*articulation rate*), sondern durch den geringeren Anteil von Pausen (bzw. Pausen, Füllern, Wiederholungen und Reparaturen) an der Sprechzeit. Dies zeigt sich an den Ergebnissen zur *speech rate* bzw. zur *pruned speech rate*. Längere Pausen (bzw. weniger produktive Sprechzeit) deuten auf mehr Planungsbedarf bei der KG hin. Die KT der IG scheinen dagegen vermehrt auf fertige Muster zurückgreifen zu können, was auf vorgängige Prozeduralisierung der gelernten Sprache (de Jong & Perfetti, 2011) bzw. auf das Vorhandensein von *constructions* hinweist, die abgerufen werden können. Die Ergebnisse zur schriftlichen Korrektheit, die keine Unterschiede zwischen IG und KG zeigen, widersprechen dieser Interpretation nicht. Möglicherweise manifestierte sich wegen des relativen Zeitdrucks in der *Elicited Imitation* die „ballistische“ Natur von automatisierter Sprache (Segalowitz & Hulstijn, 2009), während in der schriftlichen Prüfung Zeit blieb, anderes Wissen einzubeziehen, das korrektiv wirkte. Offen bleibt, um welche Formen des Wissens es sich dabei handelte: Wurden verschiedene erworbene Chunks verglichen oder wurde abstraktes (Regel-) Wissen beigezogen, das die KT selbst aus ihren Spracherfahrungen abgeleitet hatten?

Die Ergebnisse in den mündlichen Testteilen sollten auch aus der Perspektive der Forschungsergebnisse zur Performanz bei Sprechaufgaben betrachtet werden, besonders unter Berücksichtigung der *Tradeoff*-Hypothese (Skehan, 2009). Ausgangspunkt dieser Hypothese ist die allgemein anerkannte Tatsache, dass unsere kognitiven Kapazitäten beschränkt sind: Die Kapazität des Arbeitsgedächtnisses ist begrenzt, und wir können nur eingeschränkt aufmerksam sein. Beim Sprechen besteht deshalb zwischen den Ansprüchen nach sprachlicher Komplexität, Korrektheit und Flüssigkeit²² eine Konkurrenzsituation.

Bei unseren stark vorstrukturierten Testaufgaben war die Komplexität vermutlich kein wichtiger Faktor und zudem für beide Gruppen grundsätzlich gleich. Die kognitiven Ressourcen konnten deshalb im Prinzip weitgehend auf sprachliche Korrektheit und Flüssigkeit verwandt werden. Es scheint nun so, dass die IG der Flüssigkeit den Vorzug gab

22 Das Verhältnis von *complexity*, *accuracy* und *fluency* (kurz: CAF) ist Gegenstand einer andauernden Debatte in der angewandten Linguistik und in der Sprachtestforschung.

und die KG der Korrektheit. Aus der Sicht der Vertreter eines lexikalischen Ansatzes (Nation & Newton, 2009), müsste das nicht so sein. Es besteht vielmehr die Erwartung, dass wegen der als ganze Elemente zur Verfügung stehenden Chunks nicht nur die Flüssigkeit, sondern auch die Korrektheit gesteigert würde. Für den Fall, dass vorhandene Chunks noch grammatisch angepasst werden müssten, könnte wiederum erwartet werden, dass schneller und leichter abrufbare Chunks kognitive Ressourcen freisetzen, um den Fokus auf diese Anpassungen bzw. Korrekturen zu richten. Beides zeigt sich so in den Daten nicht. Denkbar ist, dass sich im vorliegenden Fall nicht die beschränkten kognitiven Ressourcen während der Sprachproduktion in den Ergebnissen manifestieren, sondern vielmehr der unterschiedliche Charakter des sprachlichen Wissens und Könnens, das sich im Unterricht der beiden Gruppen entwickelt hat. Die IG verfügt vermutlich über mehr automatisierte, aber fehlerbehaftete Chunks, die mit „ballistischer“ Kraft aus dem Gedächtnis hervorschiessen, während die KG über die Sprachkenntnisse und die notwendige Aufmerksamkeit verfügt, um – auf Kosten der Sprechflüssigkeit allerdings – mündliches Deutsch etwas korrekter zu produzieren.

Auch unsere Ergebnisse sprechen gegen die *Non-interface*-Position (Krashen, 1982): Der Erwerb von frei verfügbarer Grammatik für den eher spontanen Gebrauch scheint gerade *durch* das explizite Lernen besser gelungen. Allerdings zeitigt der Unterricht der IG, in dem die zu erwerbenden Grammatikphänomene lediglich systematisch gehäuft auftreten, im Form-fokussierten schriftlichen Sprachgebrauch einen vergleichbaren Lernzuwachs. Beiläufiges (inzidentelles) Lernen fand offensichtlich statt und stellt Grammatikressourcen zur Verfügung, auf die bei schriftlichen Grammatikitems zurückgegriffen werden kann.

Wie steht es aber um die Vermutung, dass die auf der Basis von Chunks unterrichteten Gruppen zu gegebener Zeit aufgrund der angesammelten Chunks die grammatische Analyse der bereits verwendeten Sprache nachvollziehen werden? – Das ist grundsätzlich denkbar. Es hängt davon ab, was der beobachtete Unterschied zwischen mündlicher und schriftlicher Korrektheit bei der IG aus kognitiver Sicht bedeutet. Ist der relativ gesehen grössere Lernerfolg im schriftlichen Gebrauch ein Zeichen dafür, dass auch korrektere Formen gespeichert wären, aber (noch) weniger automatisch abgerufen werden, d. h. noch weniger ballistische Kraft haben? In diesem Fall könnte es v. a. bei weiterem Unterrichtsbesuch und ganz besonders bei einem Unterricht, für den das Verfügbarmachen von formal korrekter Sprache ein Ziel ist, durchaus der Fall sein, dass eine erfolgreiche Analyse der sprachlichen Formen stattfindet.

Die Tatsache, dass sich nur bei den Lernenden der IG überhaupt eine Entwicklung bei der Flüssigkeit gezeigt hat (anhand der beiden *speech rates*), ist ein wichtiger Be-

fund. Sie bestätigt eine Erwartung von Lewis (1993) und anderen Vertretern des Lexikalischen Ansatzes. Flüssigkeit ist einer der zentralen Bereiche, in denen sich Sprachkompetenz entwickeln muss. Bei der KG scheint es da einen Stillstand zu geben. Offen ist, wie weit dieser lediglich ein Symptom konkurrierender Ansprüche von Korrektheit und Flüssigkeit um die limitierten Gedächtniskapazitäten ist. Es könnte sich durchaus so verhalten, denn v. a. das Ergebnis bei der *pruned speech rate*, die auch Aspekte der Planung und der Selbstkorrektur widerspiegelt, lässt vermuten, dass in der KG beim Sprechen tatsächlich mehr „gearbeitet“ wurde als in der IG. Ein Unterrichtsziel dieser Gruppe müsste in Zukunft darin bestehen, durch Wiederholungsaktivitäten vermehrt automatisierte Sprache (abrufbare Chunks bzw. *constructions*) zu entwickeln. Sicher ist, dass Sprachkontakte bei flüssigerem Sprachgebrauch leichter fallen. Dies wiederum bietet längerfristig mehr Chancen, auch ausserhalb des Kurses weiterzulernen.

7 Fazit

Die praktische Frage, die am Ursprung dieser Studie stand, muss aufgrund der Ergebnisse differenziert beantwortet werden. Ein chunkorientierter Unterricht, der auf explizite Grammatikarbeit (*focus on forms* und *focus on form*,²³ Long, 1991) verzichtet, lässt sich zwar in niederschweligen Sprachkursen erfolgreich umsetzen, stösst aber bei Lernenden mit Vorerfahrungen im Fremdsprachenunterricht und solchen mit Zielen im Bereich der Korrektheit auf Vorbehalte, während Schulungsgewohnte gemäss Einschätzung der Kursleitenden dieses Konzept als Erleichterung empfinden. Die Kursleitenden selbst nennen einerseits eine wichtige Voraussetzung für das Gelingen des Unterrichts nach dem Interventionsprinzip, nämlich die, dass angepasste Kursmaterialien vorliegen müssen, und andererseits einen wichtigen problematischen Punkt: das im vorliegenden Interventionskonzept offenbar unvermeidliche, häufige Wiederholen von nicht-zielsprachlichem Deutsch. Genau dieses Merkmal wiederum kann dafür verantwortlich sein, dass die Lernenden im Bereich der mündlichen Korrektheit weniger Fortschritte aufweisen als die Lernenden der KG, die nach einem herkömmlichen Konzept mit zahlreichen Aktivitäten gelernt hat, die auf Sprachformen fokussieren. Umgekehrt hat die KG, im Unterschied zur IG, im ebenfalls wichtigen Bereich der mündlichen Flüssigkeit während der Interventionsphase keine messbaren Fortschritte gemacht. Diese Sachlage spricht für *balanced learning opportunities* (Ellis, 2005b). Im Unterrichtskonzept der IG braucht es zumindest eine stärkere Fokussierung auf die Korrektheit der Chunks, die automatisiert werden. Das bedeutet nicht zwingend systematische Grammatikarbeit, sondern insgesamt einen stärkeren Fokus auf sprachliche Formen, vielleicht einfach ein stärkeres Beharren auf zielsprachlich korrekten Formen in der Sprache, die von den Lernenden beim Flüssigkeitstraining verwendet wird. Im Konzept der KG fehlt dagegen die gezielte Förderung der Automatisierung durch häufige Wiederholungsaktivitäten.

Aus Forschungssicht ist eine vertiefte Analyse der vorliegenden mündlichen Produktionen wünschenswert, anhand derer die Flüssigkeitsindikatoren ermittelt wurden, um auf diese Weise gegebenenfalls mehr über Unterschiede in der „Formulierungsarbeit“ und damit in der kognitiven Basis für die Produktionen der beiden Gruppen zu erfahren. Gibt es beispielsweise je nach Gruppe spezifische Unterschiede in der Art der formalen Fehlleistungen? Gibt es Unterschiede hinsichtlich der Art der gefüllten Pausen, der Reparaturen und Wiederholungen? Weiter würde es sich lohnen, den Einfluss eines gezielten Flüssigkeitstrainings im Rahmen eines klassischen schulischen Fremdsprachen-

23 *Focus on forms*: systematischer expliziter Grammatikunterricht; *focus on form*: expliziter Grammatikunterricht auf der Basis der Klärungsbedürfnisse, die sich aus einem kommunikationsorientierten Unterricht ergeben.

unterrichts zu erforschen, wo in der Regel weder intensiv unterrichtet wird, noch viele ausserschulische Sprachkontakte bestehen. Dabei sollten unterschiedliche Kompetenzstufen einbezogen werden, um dem Umstand Rechnung zu tragen, dass möglicherweise die Grösse des Repertoires bereits vor dem Training verfügbarer korrekter Formen zu unterschiedlichen Ergebnissen in der Entwicklung der formalen Korrektheit führen kann.

Fluidité et correction – fluidité ou correction ?

Synthèse des résultats d'une étude expérimentale portant sur des cours intensifs à bas seuil d'allemand langue seconde, destinés à des adultes débutants

—
Peter Lenz, Malgorzata Barras, Fabienne Manz¹

¹ Un immense merci au Prof. Thomas Studer, responsable du domaine de recherche, pour ses précieux conseils concernant la réalisation de ce projet et l'organisation du présent rapport.

Résumé

Cette étude examine l'efficacité d'un enseignement basé principalement sur l'approche lexicale (*lexical approach*) (Lewis, 1993 ; Lewis, 1997), soit sur l'apprentissage de séquences de mots avec des parties variables (*chunks*) et sur un entraînement régulier à la fluidité (Nation & Meara, 2002), dans le cadre de cours d'allemand langue seconde, destinés à des adultes débutants en partie peu scolarisés. L'étude, conçue de manière quasi-expérimentale, a comporté un groupe expérimental (GE) et un groupe témoin (GT). Elle a porté sur douze classes, réparties en six paires de deux classes, qui ont étudié un manuel de cours complet. Pour chaque paire, une classe a suivi un enseignement basé sur le concept expérimental et utilisé les ressources didactiques adaptées, tandis que l'autre classe a travaillé selon le concept et avec les ressources didactiques témoin qui, pour l'essentiel, correspondaient au matériel didactique original. Le concept témoin repose principalement sur un enseignement explicite de la grammaire et sur une étude du vocabulaire par mots isolés. Lors d'un test initial puis d'un test final, diverses compétences orales et écrites des participants ont été évaluées. En outre, des entretiens menés avec les enseignants ont permis de recueillir des informations quant à l'adéquation des concepts didactiques aux besoins du groupe cible ainsi qu'aux particularités des processus d'enseignement et d'apprentissage propres aux deux groupes.

Les entretiens donnent à penser que, dans l'ensemble, les deux concepts permettent un enseignement de bonne qualité. Les participants du groupe expérimental qui ont de l'expérience dans l'apprentissage des langues étrangères et ceux qui mettent la priorité sur la correction de leur expression déplorent parfois que les ressources et l'enseignement ne soient pas plus axés sur la grammaire. Ils voient généralement dans l'entraînement à la fluidité une bonne occasion d'apprendre. Les observations des enseignants confirment que le concept expérimental correspond mieux à des apprenants peu scolarisés, alors que le concept témoin est plus adapté à des apprenants disposant d'un bon bagage scolaire. Les enseignants du groupe expérimental regrettent parfois l'absence de la métalangue grammaticale pour corriger les erreurs. De plus, ils déplorent que des erreurs de langue soient souvent répétées et assimilées lors de l'entraînement à la fluidité, et qu'ils n'aient pas suffisamment d'occasions de les corriger.

Les résultats de l'évaluation des compétences montrent une progression entre les tests initiaux et finaux dans presque tous les domaines. Seules les mesures de fluidité font exception. Les deux principaux indicateurs de fluidité n'affichent aucune progression pour le groupe témoin alors que le groupe expérimental progresse. Les progrès en vitesse d'élocution (*speech rate*, ou le nombre de syllabes par seconde, toutes pauses comprises) et en vitesse d'élocution corrigée (*pruned speech rate*, ou le nombre de syllabes par seconde à l'inclusion des pauses silencieuses, mais à l'exclusion des pauses

remplies, des répétitions et des autocorrections) indiquent clairement que la tâche de production orale demande moins de planification et de contrôle aux participants du groupe expérimental qu'à ceux du groupe témoin. Les résultats détaillés d'une autre partie du test oral suggèrent que ce gain en fluidité s'accompagne d'une perte en correction. Toutefois, du point de vue de la fonction et du contenu, les compétences orales des deux groupes n'affichent pas de différence. Par ailleurs, aucune différence de niveau n'apparaît entre les deux groupes concernant la partie écrite du test, pas même dans le domaine de la morphosyntaxe, où l'on s'attendait pourtant à ce que le groupe témoin obtienne de meilleurs résultats. En conclusion, cette étude analyse les résultats obtenus à la lumière des approches théoriques relatives à l'acquisition et la production langagières. Elle présente également ses conclusions quant aux concepts didactiques comparés et aux futures recherches.

1 Introduction

Le projet « Impact des procédés didactiques sur l'apprentissage des migrants adultes dans les cours de langue seconde », qui se trouve être à la base de la présente étude, a été réalisé dans le cadre du programme de recherche 2011-2015 du Centre scientifique de compétence sur le plurilinguisme, avec le soutien du Secrétariat d'État aux migrations (SEM). Il visait à apporter, à partir d'observations empiriques, un éclairage nouveau sur une question soulevée par les experts au cours de l'élaboration du « Curriculum-cadre suisse pour la promotion linguistique des migrants » (Lenz, Andrey & Lindt-Bangerter, 2009). Il s'agissait en effet de savoir si les cours d'allemand langue seconde, s'adressant à des participants peu scolarisés, doivent s'appuyer sur une approche didactique conventionnelle, dont les fondements sont un enseignement explicite et analytique de la grammaire (*focus on forms*, selon Long, 1991) et une étude du vocabulaire par mots isolés, ou si ce type de cours appelle plutôt un apprentissage et une automatisation ciblés de séquences de mots avec des parties variables (*chunks* en anglais). La présente étude tente d'apporter des éléments de réponse fondés sur des observations concrètes et objectives.

L'étude a été menée selon un modèle quasi-expérimental, avec un groupe expérimental et un groupe témoin. Six paires de classes ont été formées. Pour chaque paire, une classe a suivi un enseignement basé sur le concept de *chunks*, soit sur l'apprentissage de séquences de mots avec des parties variables. L'autre classe a travaillé selon une approche fortement axée sur un enseignement explicite de la grammaire et sur une étude du vocabulaire par mots isolés, comme le prévoyait du reste le concept didactique original.

Au début, quatre organisations proposant des cours à bas seuil d'allemand langue seconde ont collaboré au projet. En fin de compte, l'étude a pu être conduite comme prévu avec trois des organisations, pour un total de douze classes. La quatrième classe a dû renoncer à participer au terme d'une première phase, en raison de l'hétérogénéité de ses classes, qui aurait nécessité une nouvelle répartition des participants. Grâce à un engagement fort de toutes les parties, la collaboration s'est déroulée de manière très positive.²

Les ressources didactiques ont constitué l'un des piliers de l'étude et de sa mise en œuvre. Les manuels d'allemand langue seconde, utilisés dans le cadre de l'étude, étaient ceux que les organisations proposent en temps normal. Deux des organisations

2 Nous remercions chaleureusement les organisateurs de cours suivants: aoz (Zurich), Formazione (Berne), Lernpunkt (Berne) et ECAP (Lucerne).

s'appuyaient sur le même manuel et la troisième sur un autre.³ Divers éléments des deux manuels ont dû être modifiés ou complétés de manière ciblée pour la mise en œuvre du concept centré sur les *chunks*. Les ressources en question ont finalement été réimprimées sous forme de brochures reliées.⁴

Ce type d'enquête est rare. D'une part, l'apprentissage des langues par des migrants adultes est encore peu étudié, même si la question de la promotion linguistique dans le domaine migratoire est devenue un thème de société. D'autre part, rares sont les interventions quasi-expérimentales qui s'étendent sur plusieurs semaines et touchent à la conception de cours dans leur globalité. Elles représentent un défi particulier tant sur le plan pratique que d'un point de vue scientifique, notamment parce que les facteurs extérieurs, susceptibles de diluer l'effet du concept expérimental, sont plus difficiles à contrôler sur la durée du cours que lors d'interventions ponctuelles. En revanche, les interventions s'intégrant à l'écosystème d'un cours aboutissent à des conclusions plus directement transposables dans la pratique.

3 Il s'agit des manuels d'allemand langue seconde (versions adaptées pour la Suisse) *Schritte plus 1. Niveau A1/1. Ausgabe Schweiz* (Niebisch, Penning-Hiemstra, Specht & Bovermann, 2011) éd. Hueber Verlag et *Studio d A1. Schweizer Ausgabe* (Funk, Kuhn, Demme & Bayerlein, 2005) éd. Cornelsen Verlag.

4 Nous remercions ici tous ceux qui nous ont accordé le droit d'utiliser leurs ressources didactiques.

2 Situation initiale et bases théoriques

Cette étude compare le développement des compétences linguistiques d'adultes débutants dans le cadre de cours intensifs d'allemand langue seconde, dans lesquels sont appliqués deux concepts didactiques différents.

L'objectif général de ces cours est d'améliorer la capacité d'action langagière des participants afin qu'ils puissent davantage prendre part à la vie sociale de leur pays d'accueil. Toutefois, notre étude n'évalue pas uniquement la capacité d'action communicative au travers de tâches communicatives, mais se concentre sur différentes ressources linguistiques et donc sur des composantes de la capacité d'action communicative, telles que la fluidité ou la correction grammaticale. Ces choix nous inscrivent dans la tradition de recherche de la linguistique appliquée. Ils permettent des mesures plus nuancées et plus aisément interprétables, et se justifient notamment par le fait que chez les débutants, le développement d'un répertoire linguistique servant de base à l'utilisation de la langue est prioritaire.

Nombreux sont ceux qui voient la fluidité comme le but ultime de l'apprentissage d'une langue. Le terme «fluidité» désigne la capacité à utiliser une langue rapidement, sans difficulté et correctement (Segalowitz, 2005). Du point de vue de la cognition, de la physiologie et de l'articulation, la fluidité requiert l'automatisation. L'acquisition d'un comportement plus complexe exige de l'attention, de la concentration et de l'effort. La performance est en principe lente et comporte des erreurs, mais elle s'améliore par la répétition, de manière que les actions sont finalement réalisées avec plus de facilité (Aguado, 2002). L'automatisation n'implique pas seulement l'accélération des processus langagiers, mais exige également une restructuration des processus de mémorisation (Anderson *et al.*, 2004; Gatbonton & Segalowitz, 2005), laquelle permet de soulager la mémoire de travail lors de l'utilisation du langage. L'automatisation se caractérise par un traitement de la langue plus rapide, facile, inconscient et inarrêtable (traitement «balistique» ou *unstoppable processing*), même si ces caractéristiques ne doivent pas nécessairement apparaître ensemble (Segalowitz & Hulstijn, 2009). Selon Segalowitz & Hulstijn, le traitement «balistique» est le critère déterminant de l'automatisation. On dit du traitement de la langue qu'il est «balistique» lorsqu'il est relativement imperméable aux perturbations. Dans le cadre de l'apprentissage d'une langue seconde, cette imperméabilité des automatismes pourrait être l'une des causes de la fossilisation, soit du fait que des formes linguistiques qui ne correspondent pas aux normes de la langue cible restent figées pendant longtemps, voire définitivement chez l'apprenant⁵ (Bybee, 2008).

5 Long (2005) doute de la réalité de la fossilisation (définitive) comme phénomène cognitif. Dans son analyse critique portant sur plusieurs études, il montre que cet état

Si l'automatisation est un élément central du développement du langage, se pose alors la question de comment l'acquérir, notamment dans le cadre de l'enseignement des langues étrangères.

Du point de vue de la linguistique fondée sur l'usage (*usage-based*), il s'agit d'enclencher des processus naturels au moment de l'acquisition de la langue et de les mettre à profit pour finalement parvenir à un usage automatisé de la langue. La tâche d'enseignement et d'apprentissage consiste à faire apparaître et à développer chez l'apprenant ce que l'on appelle des *constructions*. La supposition de base est que les apprenants (L1 et d'autres langues) déduisent des éléments et conventions linguistiques de leur propre expérience communicationnelle. Au terme d'un processus d'abstraction, appelé *procéduralisation*, les connaissances et compétences tirées de ces expériences sont enregistrées dans la mémoire en tant que constructions. Les constructions permettent d'associer des fonctions communicatives spécifiques à des formes linguistiques déterminées (Ellis, 2005a). Elles se composent de *chunks* cognitifs figés (*neuromotor routines*) et d'un nombre variable d'éléments interchangeable (Bybee, 2008). Le mot construction peut désigner des mots, des expressions ou des phrases entières. L'expression semi-figée «J'aimerais bien...» est un exemple de construction qui laisse une place à un élément variable. Les constructions doivent être reconnues dans les manuels, puis automatisées par la répétition, avant de pouvoir se manifester dans l'utilisation de la langue. Le langage peut devenir fluide dès que l'apprenant dispose de suffisamment de constructions automatisées et surtout de constructions de plus grande envergure en regroupant d'autres, plus petites. Ainsi, les constructions soulagent la mémoire de travail, dont les capacités sont limitées, et libèrent des ressources cognitives qui peuvent être concentrées sur le pilotage de l'utilisation de la langue, par exemple (Skehan, 2009).

Bybee (2008) propose des réflexions sur un plan d'étude favorisant le développement de constructions. Un élément clé consiste à concevoir des conditions d'apprentissage permettant de compenser la plus faible exposition à la L2 (qu'à la L1). Il suggère ainsi de procéder dans un premier temps à une sélection de séquences de mots fondée sur une analyse de corpus. Sur cette base, il s'agit ensuite de transmettre les constructions importantes d'une langue à l'aide d'exemples typiques. Pour que les modèles soient reconnus comme tels et que l'acquisition du langage s'amorce, les appre-

de déficience, tant redouté dans l'acquisition des langues étrangères et du langage, n'existe probablement souvent pas là où il est diagnostiqué. Selon lui, les propriétés du langage qui ne seraient pas conformes aux normes indiquent avant tout que le processus d'acquisition du langage n'a pas été poursuivi, pour différentes raisons.

nants doivent être confrontés de manière répétée aux mêmes modèles et de manière ciblée à des variantes modifiées.

Segalowitz & Hulstijn (2009) attirent l'attention sur le fait qu'un cours de ce type, bien que recommandable d'un point de vue linguistique pour soutenir l'automatisation, risque de rencontrer des résistances dans la pratique. En effet, s'il fait la part belle à la répétition pour que la langue s'automatise, il se heurte à un obstacle : les personnes concernées pourraient le percevoir comme décontextualisé et peu attirant. Par contre, un cours fondé sur la communication, bien que pertinent du point de vue du contenu, prévoirait trop peu de répétitions pour véritablement soutenir l'automatisation. Pour les deux auteurs, une solution se trouve dans des concepts orientés sur des besoins communicationnels et transmettant des phrases polyvalentes en langue cible, répétées avec régularité dans le cadre d'activités privilégiant la communication et un contenu pertinent.

Dans le cadre de notre étude, la principale différence didactique entre le concept expérimental et le concept témoin concerne le traitement de la grammaire. Dans le groupe témoin, la grammaire est expliquée de manière explicite et systématique, puis elle est entraînée (*focus on forms*). Dans le groupe expérimental, par contre, la grammaire n'est pas étudiée en tant que telle. Se pose alors une question depuis longtemps débattue dans la recherche en acquisition du langage et en didactique des langues étrangères : les connaissances transmises en classe de manière explicite, règles à l'appui, puis entraînées, peuvent-elles être spontanément utilisées par l'apprenant de la même manière que les connaissances qu'ils ont eux-mêmes déduites du stimulus en langue étrangère ? La position la mieux connue sur cette question est sans doute celle de Krashen (1982) qui affirme de façon radicale que cela n'est pas possible (*non-interface position*) et que les connaissances fondées sur les règles ne servent qu'à la supervision de l'utilisation de la langue. Depuis, s'est imposée la vision selon laquelle les processus d'apprentissage complexes, tels que l'acquisition des langues, peuvent certainement passer par des connaissances conscientes (savoir déclaratif) pour devenir des compétences automatisées. Le savoir déclaratif devient alors un élément du savoir procédural – voir à ce propos le célèbre modèle du développement cognitif d'Anderson (Anderson *et al.*, 2004). Par ailleurs, une grande méta-étude, portant sur le succès de différentes formes d'enseignement (Norris & Ortega, 2001), vient appuyer l'idée qu'un cours qui se concentre systématiquement ou ponctuellement sur des phénomènes linguistiques aura plus de succès qu'une forme d'enseignement purement implicite. Selon toute vraisemblance, un enseignement développant également du savoir linguistique explicite permettra les progrès les plus importants. Un article plus récent de Spada & Tomita (2010), dont l'am-

pleur dépasse l'étude précédente puisqu'il porte sur plus de formes de test ouvertes et de résultats concernant l'enseignement implicite, confirme la supériorité d'un enseignement comprenant des parties explicites sur un enseignement purement implicite. Notre concept expérimental regroupe certes de nombreuses formes d'exercices centrés sur des phénomènes linguistiques précis, mais ces phénomènes ne sont jamais explicitement pris pour sujet d'étude. La différence avec le concept du groupe témoin reste donc claire, puisque ce dernier prévoit un travail explicite plus important sur la grammaire.

Pour la mise en œuvre du concept expérimental, nous nous appuyons en première ligne sur une approche issue de la didactique des langues étrangères ayant acquis une certaine notoriété depuis les années 1990 : l'approche lexicale (*lexical approach*) de Michael Lewis⁶ (p. ex. Lewis 1993 ; 1997 ; 2000). Nous l'avons complétée principalement par des propositions de Nation (2001) relatives à l'entraînement à la fluidité.

L'approche de Lewis repose avant tout sur des connaissances tirées de la linguistique des corpus. Il se penche sur la fréquence de liens plus ou moins figés entre des éléments de vocabulaire. Pour Lewis, la clé de la grammaire réside dans les éléments lexicaux. Il affirme ainsi que « la langue ne procède pas d'une grammaire lexicalisée, mais d'un lexique grammaticalisé » (Lewis, 1993 : 34). Il recommande de fonder l'apprentissage sur des séquences de mots qui apparaissent réellement ensemble, parce que la langue se compose en grande partie de ce type d'unités préfabriquées. Celles-ci ne peuvent pas être manipulées librement, c'est-à-dire selon les possibilités que la grammaire offrirait en soi. Lewis appelle ces unités des *chunks*. Il distingue deux catégories principales de *chunks* : les collocations et les expressions (Lewis, 1997). Par collocation, on entend les mots qui apparaissent souvent dans le voisinage immédiat l'un de l'autre (p. ex. « prendre une douche » et non pas « faire la douche » comme dans d'autres langues). Pour ce qui est des expressions, Lewis distingue encore les expressions figées (p. ex. « bonne nuit ») et les expressions semi-figées, qui se caractérisent par une ou plusieurs places tolérant des variations (p. ex. « j'aimerais bien... »).

Dans notre étude, nous qualifions autant des expressions figées que semi-figées de *chunks*, en accord avec Lewis. Nous comptons également comme expressions figées des unités fonctionnelles à un mot comme « merci ! ». Quant aux collocations, le concept expérimental ne les traite pas comme une sous-catégorie à part, mais les introduit comme éléments constitutifs des expressions.

6 Moins connus, Willis (1990) et Nattinger & DeCarrico (1992) ont également tenté de mettre en pratique les résultats de la recherche sur les *chunks* dans les plans d'études des langues.

L'enseignement d'après l'approche lexicale consiste à introduire des *chunks* de longueur variable dès les premiers stades d'apprentissage, sans toutefois les analyser grammaticalement (Lewis, 1993). Ces *chunks* sont transmis dans le cadre d'un cours axé sur la capacité d'action et sur les besoins de l'apprenant, qui doit pouvoir disposer le plus rapidement possible de moyens linguistiques lui permettant de communiquer au quotidien. Le but d'un cours basé sur les *chunks* est qu'il soutienne la fluidité, en mettant à disposition de l'apprenant des éléments plus longs auxquels il peut recourir en entier. Cet aspect favorise la réussite de la communication. L'apprentissage et l'exercice de la syntaxe, ainsi que l'apprentissage par cœur de nouveaux mots isolés choisis arbitrairement revêtent moins d'importance que dans un cours traditionnel. En revanche, la compréhension, d'abord orale puis écrite, est particulièrement entraînée. L'apprentissage actif à l'aide d'un carnet lexical personnalisé (*lexical notebook*), sorte de dictionnaire de *chunks*, et avec des cahiers d'apprentissage (Lewis, 1997) sont également privilégiés.

Dans la littérature spécialisée, les avis concernant l'approche lexicale sont partagés. De façon générale, le nombre important de *chunks* dans le langage n'y est pas remis en question. Wray & Perkins (2000) indiquent que la langue d'un locuteur natif adulte (de l'anglais) se compose à environ 70% d'expressions protocolaires, collocations incluses. On y trouve également des références à l'utilité de l'apprentissage des *chunks* pour un usage communicatif dès le début de l'étude d'une langue. Si l'apprenant étudie les *chunks* en tant qu'éléments de vocabulaire, il dispose d'un moyen de communication (correct), même s'il ne dispose pas des éléments de connaissance nécessaires à la compréhension du fonctionnement de la langue. Or, une communication améliorée grâce aux *chunks* peut engendrer de nouvelles occasions d'apprendre (Aguado, 2002 ; Nation & Newton, 2009 ; Wray, 2002). Quant au développement de connaissances grammaticales implicites et explicites, on estime qu'à partir d'un certain stade de développement, les apprenants procéderont à l'analyse de la collection d'expressions qu'ils ont acquises pour en évaluer la régularité (p. ex. Lewis, 1997 ; Wray, 2000). Les résultats de cette analyse sont ensuite enregistrés aux côtés des *chunks* originaux et forment une nouvelle ressource sur laquelle l'apprenant peut s'appuyer (Wood, 2002).

Une question décisive est de savoir si cette analyse a vraiment lieu et, le cas échéant, à quel stade. Bybee (2008) propose donc, dans la lignée de la linguistique basée sur l'usage (*usage-based linguistics*), de compléter l'approche lexicale par des éléments didactiques spécifiquement destinés à faire naître des constructions morpho-syntaxiques.

Dans notre concept didactique expérimental, l'approche lexicale est complétée par un entraînement ciblé à la fluidité : la répétition doit ainsi faciliter l'acquisition. Ce

choix s'appuie notamment sur le concept des « quatre piliers » (*four strands*) pour un bon enseignement du vocabulaire, développé par Nation (p. ex. Nation & Meara, 2002). Selon ce concept, 25% du temps de cours devrait être réservé à des activités faisant appel à des éléments lexicaux déjà connus, sans se concentrer sur des aspects linguistiques particuliers mais en introduisant une légère pression temporelle. Ainsi, les apprenants disposent ensuite d'éléments qu'ils peuvent utiliser avec fluidité. Deux variantes sont à distinguer : dans le premier cas, du lexique déjà connu est répété à plusieurs reprises, par exemple sous une pression temporelle croissante. Dans le second cas, le but visé est l'enrichissement du vocabulaire. À cette fin, du lexique déjà connu est abordé dans de nouvelles combinaisons et de nouveaux usages, p. ex en lisant des textes plus longs mais simplifiés. Dans le concept expérimental, seule la première variante est mise en œuvre.

3 Objectifs de l'étude et questions de recherche

Un des objectifs principaux de l'étude était le développement d'un concept d'enseignement valable pour le groupe expérimental, puis sa mise en pratique par l'adaptation de ressources didactiques. Il s'agissait en outre de trouver des organisateurs de cours de langue disposés à travailler pendant deux à quatre mois en suivant nos directives. La question centrale concernait les progrès des participants aux cours, qui devaient être mesurés à l'aide d'instruments adaptés, puis analysés.

L'étude visait à répondre aux questions suivantes :

- i. Un concept didactique conçu selon les fondamentaux du concept expérimental ainsi que les matériaux didactiques correspondants peuvent-ils être élaborés conjointement et mis en œuvre avec succès dans des cours à bas seuil d'allemand langue seconde ?
- ii. Le concept expérimental est-il accueilli avec intérêt et accepté par les enseignants et les participants ?
- iii. Le concept expérimental est-il adapté aux besoins notamment des apprenants peu scolarisés ?
- iv. La progression des participants du cours expérimental et celle des participants au cours témoin est-elle la même dans les différents sous-domaines de compétences langagières ?

Les réponses aux questions i à iii devaient s'appuyer sur des entretiens menés avec les enseignants.

La progression des participants a été mesurée à l'aide de tests oraux et écrits. Quant à la comparaison des deux groupes, elle se basait sur l'hypothèse selon laquelle le concept expérimental serait considéré comme plus efficace que le concept témoin si, par rapport au groupe témoin, le groupe expérimental atteignait les résultats suivants :

- iv.a. Utilisation plus fréquente de *chunks* en langue cible
- iv.b. Expression plus fluide
- iv.c. Plus grande correction de l'expression orale (correction intuitive) ; correction équivalente ou très proche à l'écrit (correction en application de règles)
- iv.d. Connaissances lexicales comparables du point de vue sémantique et capacité comparable à communiquer des contenus sans prise en compte de la correction formelle

4 Méthode

4.1 Conception de l'étude

Nous avons choisi de mener l'étude de manière quasi-expérimentale, avec un groupe expérimental et un groupe témoin. Nous avons exclusivement travaillé dans le cadre de cours intensifs à bas seuil,⁷ où les participants consacraient à l'apprentissage de l'allemand quatre à cinq jours par semaine, soit 8 à 11h de cours hebdomadaires. Suivant les organisateurs de cours, entre huit et seize semaines étaient prévues pour parcourir un volume du manuel d'allemand.

Afin de compenser les effets de ces écarts de durée et d'éventuelles autres différences spécifiques à chaque organisateur, des paires de classes parallèles, regroupant des débutants,⁸ ont été composées avant le début d'un nouveau cours. Pour chaque paire, une classe a suivi le concept expérimental tandis que dans l'autre a été appliqué le concept témoin. Les participants ont été attribués au groupe expérimental ou au groupe témoin selon le principe de l'échantillonnage stratifié. La connaissance que les organisateurs avaient des participants a été mise à profit pour qu'entre les deux classes de chaque paire, les caractéristiques pouvant potentiellement influencer l'apprentissage comme l'origine, la langue maternelle et le bagage scolaire, soient équilibrées.

Les organisateurs ont ensuite attribué des enseignants aux classes ainsi formées. En tout, sept paires de classes ont été constituées de cette manière sur les quatre organisations participantes. Par la suite, une des paires a dû se retirer de l'étude parce qu'en raison de nos principes de répartition, la composition des classes était trop hétérogène, ce qui a mené à des frustrations et a nécessité une répartition par niveaux.

Notre intervention s'est déroulée en deux phases de durée égale. Pendant la première phase, qui revêtait un caractère préliminaire, nombre des participants ont acquis leurs premières connaissances d'allemand, alors que les enseignants s'habituèrent à travailler avec nos concepts et ressources didactiques. Au terme de cette première phase, nous avons réalisé un test initial, dont la structure était pratiquement identique à celle du test final. Certains participants ont quitté le cours après la première partie, alors que d'autres l'ont rejoint. Tous les nouveaux venus ont également passé le test

7 L'accès aux cours pour débutants dits « à bas seuil » est facilité à plusieurs égards et particulièrement en ce qui concerne les prérequis scolaires et linguistiques ainsi que le financement.

8 Cette désignation est utilisée par les organisateurs de cours pour parler des personnes intéressées à un cours de niveau A1.1 d'après le CECR (Conseil de l'Europe & Goethe-Institut Inter Nationes, etc., 2011).

initial. En règle générale, le test final a été passé au cours de la dernière semaine de la deuxième moitié du cours. L'analyse quantitative reflète donc les résultats de la deuxième moitié du cours.

4.2 Intervention didactique

L'intervention menée dans le cadre de cette étude a touché divers aspects de la conception du cours. Elle a préalablement été testée de manière approfondie dans le cadre de cours intensifs, puis retravaillée et mise en œuvre au moyen d'ateliers d'introduction avec les organisateurs et les enseignants, de visites de classes, d'entretiens, de mémentos pour l'enseignement, de journaux d'enseignement et surtout de ressources didactiques, lesquelles ont joué un rôle central. L'intervention a concerné le groupe expérimental et, dans une moindre mesure, le groupe témoin.

L'enseignement dans les deux groupes d'une même paire de classes est comparable à plusieurs égards. Premièrement, les deux cours reposent sur le même manuel (les thèmes, par exemple, sont les mêmes). Deuxièmement, le nombre d'heures et l'infrastructure à disposition sont identiques et les enseignants des deux classes se coordonnent pour la durée des devoirs. Toutefois, eu égard au concept appliqué, l'accent est placé systématiquement sur des éléments différents et d'autres ressources didactiques sont utilisées.

Les exigences découlant du concept s'appliquent également lorsque les enseignants proposent eux-mêmes des activités. Globalement, l'équipe de projet s'est efforcée de fournir suffisamment de ressources didactiques conformes au concept pour éviter que trop d'approches soient mélangées.

Les deux manuels utilisés par le groupe expérimental ont été adaptés et réimprimés. Des éléments ont été supprimés, modifiés ou complétés de manière à ce que le travail sur les *chunks* devienne un élément dominant. Les modifications apportées au concept et au manuel peuvent être résumées comme suit :

- Introduction d'un entraînement à la fluidité quotidien d'une durée de 20-30 minutes. Les éléments que comporte cet entraînement ont été pour moitié appris précédemment et pour moitié appris la veille. L'essentiel est que les ressources linguistiques utilisées soient déjà connues et aient déjà été exercées en cours. Ainsi, il s'agit souvent simplement d'activités, tirées du manuel ou des ressources supplémentaires, que l'on répète sous une

certaine pression temporelle⁹ ou du moins en sachant que la fluidité est le but de l'exercice. De par sa polyvalence, le *pointing* est une forme d'exercice particulièrement adaptée : les apprenants sont invités à s'exprimer à partir d'un mot-clé, d'une carte de vocabulaire ou d'une image. L'entraînement à la fluidité comporte également des activités écrites, comme des concours en groupe, où il s'agit d'écrire le plus possible de phrases à propos d'une image, par exemple.

- Remplacement des listes de vocabulaire et des résumés grammaticaux par un carnet de notes lexical personnalisé (*lexical notebook*), selon les recommandations de Lewis (1997). Ce carnet est rédigé à mesure que le cours avance et contient, pour chaque participant, de nombreuses phrases personnalisées.
- Complètement du manuel par des dialogues de la vie quotidienne (en raison de la quantité de matière, seul un des deux manuels est concerné). Le groupe témoin travaille également sur des dialogues de la vie quotidienne, mais selon une méthode didactique différente. Le groupe expérimental traite notamment les tournures (*turns*) et les *chunks* de manière plus approfondie. Les dialogues sont schématisés, reproduits sous forme de textes à trous et mis à disposition en tant que fiches de travail.
- Adaptation systématique des exercices. Alors que le groupe témoin travaille avec des mots ou des morphèmes, le groupe expérimental travaille avec des *chunks*. Cette adaptation repose sur les types d'exercices suivants, recommandés par Lewis (1997) :
 - Ranger (p. ex. construire des *chunks* d'utilité communicative en assemblant des blocs)
 - Classer (p. ex. remettre dans l'ordre des répliques de différents dialogues mélangés)
 - compléter (p. ex. faire des dictées à trous en complétant des *chunks* entiers ou faisant émerger des *chunks* importants). Ce type de tâche est étendu aux dialogues de la vie quotidienne, dans lesquels des phrases entières sont à compléter.

⁹ Déjà lors de la phase de test du projet, il était ressorti que la pression temporelle génère trop de stress et est mal perçue pour certaines activités, comme les dialogues. C'est pourquoi, aucune activité mettant la pression temporelle au premier plan n'a été introduite, comme la répétition du même discours en un temps toujours plus bref, recommandée par Nation.

- Remplacement des instruments de contrôle des objectifs d'apprentissage, généralement fortement centrés sur la grammaire, par des instruments analogues, centrés sur les *chunks*.

Nombre d'éléments caractéristiques des exercices apparaissant dans les manuels originaux, et donc utilisés dans les cours du groupe témoin, ont été éliminés et remplacés dans les ressources du groupe expérimental. Il s'agit avant tout des activités d'analyse des règles grammaticales, des activités d'identification et de classification de phénomènes grammaticaux, et des exercices d'application de règles morphologiques et syntaxiques ou d'adaptation grammaticale.

Seules quelques ressources complémentaires sont venues s'ajouter au matériel didactique du groupe témoin. Les enseignants se sont engagés à donner un cours d'allemand langue seconde le plus proche possible de l'esprit du manuel, tel qu'interprété dans le cadre du concept témoin. Cela impliquait notamment de limiter leurs apports (idées et ressources) personnels. Ceci a permis de définir plus clairement le mode d'enseignement dans le groupe témoin et d'uniformiser les cours entre les différents enseignants.

4.3 Récolte des données

Il convient de distinguer trois types de données récoltées : 1) les données personnelles des participants provenant des informations fournies par les organisateurs de cours ; 2) les tests oraux et écrits réalisés sur deux périodes, soit au début et à la fin de la seconde moitié du cours ; 3) les entretiens avec les enseignants.

Données personnelles des participants

Tous les organisateurs de cours obtiennent les noms et prénoms des apprenants, ainsi qu'un certain nombre de données concernant le parcours de l'apprenant. Ces dernières sont d'intérêt soit pour des raisons administratives, soit pour procéder à l'attribution à une classe. Pour constituer les groupes dans cette étude, nous nous sommes penchés en premier lieu sur les profils linguistico-culturels et scolaires. D'autres variables, qui auraient pu expliquer des divergences dans les progrès au moment de l'analyse ou être pris en compte au moment de la constitution des groupes, n'étaient que rarement à disposition, comme la perception qu'avaient les enseignants précédents de la capacité d'apprentissage de chaque participant.

Tests

Dans toutes les parties du test, le principe était d'évaluer uniquement les connaissances et capacités acquises dans le cadre du cours, les influences de l'environnement privé devant être réduites au minimum. Une version du test a été conçue pour chacun des deux manuels, avec toutefois une majorité de questions communes. Le test final se composait de questions déjà présentes dans le test initial et de questions sur le second bloc de cours. Il fallait que les tests se recoupent pour que tous les résultats de chaque domaine de compétence testé puissent être représentés sur une même échelle.

Les tests ont été menés sur place par des responsables formés. Le test écrit était passé en groupe et durait environ 80 minutes. Le test oral était passé individuellement sur ordinateur, sous la supervision d'un responsable. Le premier test oral durait 30-40 minutes, le second 35-45 minutes.

Le tableau ci-dessous offre une vue d'ensemble des domaines de compétence testés et des instruments utilisés lors des deux tests.

Domaine testé	Type de tâche [aspect encodé]
Test écrit	
1a) Compétence langagière générale à l'écrit (du point de vue du fond)	C-test ¹⁰ avec 4 textes pour 15 trous [reconstruction reconnaissable du mot original]
1b) Compétence langagière générale à l'écrit (du point de vue du fond et de la forme)	C-test avec 4 textes pour 15 trous [reconstruction complète du mot original]
2) Grammaire : morphosyntaxe (potentiellement dérivée de règles)	Construire des phrases à partir d'éléments Modifier des phrases
Grammaire : morphologie (potentiellement dérivée de règles)	Indiquer le pluriel
Lexicogrammaire : prépositions (potentiellement dérivées de règles)	Compléter des phrases
Test oral	
3a) Compréhension orale et production orale (du point de vue du fond)	<i>Elicited imitation</i> ¹¹ [nombre d'éléments de contenus restitués] <i>Elicited imitation</i> [nombre de radicaux restitués]
3b) Compréhension orale et production orale (du point de vue du fond et de la forme)	<i>Elicited imitation</i> [nombre de mots complètement restitués] <i>Elicited imitation</i> [nombre de <i>chunks</i> restitués complètement ou dont le radical est restitué]
4) Caractère idiomatique de l'expression orale (<i>chunking</i>)	<i>Elicited imitation</i> [<i>chunks</i> complets et exacts]
5) Grammaire : morphologie (prioritairement intuitive)	<i>Elicited imitation</i> [correction de la morphosyntaxe]
6) Fluidité (fluidité cognitive)	Production orale dirigée; direction par des exemples (images et enregistrements) [a. vitesse d'élocution ¹² – nombre de syllabes par seconde, toutes pauses comprises ¹³] [b. vitesse d'élocution corrigée – nombre de syllabes par seconde tous silences compris, mais toutes pauses remplies, répétitions et autocorrections exclues] [c. vitesse d'articulation – nombre de syllabes par seconde, tous silences exclus]

Tableau 1: Vue d'ensemble des domaines de compétences testés et des instruments utilisés lors du test initial et du test final

Entretien avec les enseignants

Au terme de la phase expérimentale, la collaboration avec les enseignants a été conclue par un entretien. Les entretiens avec les enseignants du groupe témoin se distinguaient de ceux du groupe expérimental uniquement sur les aspects relatifs au concept d'enseignement. Deux domaines thématiques dominants ont été abordés: 1) le concept et 2) les participants. Les entretiens ont duré entre 30 et 100 minutes.

Les seize entretiens ont été transcrits et leur contenu analysé (Mayring, 2008) afin d'en dégager les éléments pertinents pour les problématiques traitées dans le cadre de l'étude.

- i. Mise en œuvre du concept
 - a) Comparaison avec l'enseignement habituel
 - b) Mise en œuvre concrète du concept
 - c) Facteurs d'influence
- ii. Travail avec les participants et la classe
- iii. Avantages et inconvénients du concept
 - a) Point de vue de l'enseignant
 - b) Point de vue des participants (selon la perception de l'enseignant)

Traitement des résultats des tests

Les performances écrites et orales des participants ont été traitées par un tandem-encodeur dans le respect de directives écrites, détaillées pour l'analyse de données.

L'évaluation des indicateurs de fluidité reposant sur des monologues a été une tâche complexe. Les enregistrements ont été segmentés, catégorisés et comptés à l'aide du logiciel Praat (Boersma & Weenik, o.J.) et à la main, ce qui a permis ensuite de calculer les indices.

- 10 Le *C-test* est un format de texte à trous où dès la deuxième phrase, la seconde moitié d'un mot sur deux doit être complétée (C-Test-Team, o. J.).
- 11 Nous avons transposé les tâches d'*elicited-imitation* (Erlam, 2009) par ordinateur à l'aide du logiciel *PsychoPy* (Peirce, 2009). Les participants entendaient un mot ou une suite de mots, puis devaient lire un chiffre à l'écran. Après un signal, ils devaient répéter ce qu'ils avaient entendu. Le chiffre fait office de perturbateur, empêchant que les mots entendus puissent être recopiés directement à partir de la mémoire sensorielle, sans être traités (Shiffrin, 2003).
- 12 La vitesse d'élocution et les deux autres mesures sont des indicateurs de la fluidité de l'énonciation orale (*utterance fluency*) (Segalowitz, 2010). La vitesse d'élocution réunit la vitesse pure et la longueur des pauses; cette mesure technique de la fluidité se recoupe donc bien avec la fluidité perçue, contrairement à la simple vitesse d'articulation (Kormos & Dénes, 2004; de Jong, Steinel, Florijn, Schoonen & Hulstijn, 2012).
- 13 Nous avons considéré comme «silence» toute période de plus de 250 millisecondes sans articulation.

5 Résultats

5.1 Résultats des tests

Les résultats des tests écrits et du test d'*elicited imitation* ont été analysés selon le modèle de Rasch.¹⁴ Ceci a permis de transposer les résultats des 4 typologies de tests (deux manuels et deux périodes de temps) sur une même échelle et ainsi de les comparer.¹⁵ Quant aux données relatives à la fluidité de l'expression orale, elles ont pu être utilisées telles quelles, puisqu'elles renvoyaient toutes à la même tâche de production orale.

L'évaluation statistique repose sur des analyses de la covariance (ANCOVA)¹⁶ à un facteur (appartenance au groupe témoin ou au groupe expérimental) avec répétition de la mesure aux périodes t1 et t2 (test initial et test final), et est réalisée sur la base de modèles mixtes (*mixed models*).

Le nombre de participants (N) pris en considération au moment de l'évaluation varie légèrement selon les parties du test. Des 153 participants qui ont entamé le deuxième bloc de cours, seuls 129 candidats ont passé toutes les parties du test et 141 ont rendu au moins une partie du test. Aucun lien significatif n'a été observé entre les résultats au test d'entrée et la renonciation à la seconde partie du cours.¹⁷ Le tableau 2 offre une vue d'ensemble des données. Les données manquantes s'expliquent avant tout par des absences partielles de participants pour des raisons organisationnelles.

14 Les analyses statistiques ont été effectuées à l'aide de paquets du logiciel R (Chambers, 2008; Bates, Mächler, Bolker & Walker, 2015; Kiefer, Robitzsch & Wu, 2015; Lenth, 2015; Halekoh & Højsgaard, 2014).

15 Les résultats des personnes ont été rendus sous forme d'estimations pondérées des réponses les plus vraisemblables (*Weighted Likelihood Estimates*, WLE, Warm, 1989). Les échelles WLE exagèrent quelque peu la variabilité effective dans les données. Par souci de simplicité, nous avons tout de même traité les WLE comme des valeurs observées.

16 Contrairement à l'analyse de la variance, l'analyse de la covariance prend également en compte des variables explicatives dont on sait qu'elles peuvent avoir une influence sur les différences individuelles. Ici, il s'agissait de l'origine, de la durée de la résidence en Suisse, du parcours scolaire et du sexe. À ces critères vient s'ajouter la variable « paire de classes », qui tient compte du fait que les organisateurs de cours répartissaient un grand groupe de personnes intéressées en une classe témoin et une classe expérimentale. Les deux classes apprenaient ensuite simultanément et dans des conditions extérieures très similaires.

17 Ceci a été vérifié à l'aide de corrélations de Pearson entre la variable « renonciation au cours – oui ou non » et les résultats des tests au moment t1.

	GE	GT	Total
Ont débuté le bloc 2	73	80	153
Ont terminé le bloc 2 (y compris min. 1 partie de test)	66	75	141
Ratio hommes/femmes à la fin du bloc 2	34/32	34/41	68/73
Nombre d'années de scolarisation : moyenne (min. – max.)	8.8 (0-17)	9.3 (0-16)	9.1 (0-17)
Âge : moyenne (min. – max.)	33.2 (16-51)	32.3 (17-57)	32.7 (16-57)
Ont passé l'intégralité des tests au moment t2 ; données utilisables	63	66	129
Ont passé la partie écrite du test au temps t2	66	73	139
Ont passé le test oral I au temps t2 (<i>elicited imitation</i>)	63	68	131
Ont passé le test oral II au temps t2 (production orale/fluidité)	64	71	135

Tableau 2 : Vue d'ensemble des participants et des données issues des tests

Le tableau 3 offre une vue d'ensemble des résultats des tests. Les résultats des trois parties du test sont répartis sur les onze domaines examinés. Le tableau indique les moyennes par groupe et par période, ainsi que l'erreur type, entre parenthèses. Plusieurs tests de signification vérifient que les résultats peuvent être généralisés d'un point de vue statistique. Pour chaque domaine examiné, on teste a) le progrès entre t1 et t2 dans les deux groupes et b) la différence de progrès entre les deux groupes. Les données issues de cette comparaison figurent dans la dernière colonne du tableau, mise en évidence en gris, et sont d'une importance centrale pour cette étude.¹⁸

18 Pour les parties orales du test, deux valeurs p différentes sont données. La première (p) ne tient pas compte du fait que les mêmes performances sont codées et évaluées plusieurs fois pour examiner des phénomènes différents, tandis que la seconde (p_H , H pour corrigé selon Holm) tient compte de l'intervalle de confiance global habituel de 95% pour chaque partie du test. Les valeurs de p corrigées selon Holm sont donc rarement significatives.

Domaine examiné	Groupe expérimental (GE)			Groupe témoin (GT)			Comparaison GE et GT
	Test initial (t1)	Test final (t2)	Signification du progrès ¹⁹	Test initial (t1)	Test final (t2)	Signification du progrès	Différence de progrès ; Signification ²⁰
Partie écrite N = 139, n (GE) = 66, n (GT) = 73							
1a) Compétence langagière générale à l'écrit: C-test écrit / reconstruction complète	0.064 (0.126)	0.434 (0.126)	t = 5.13 p < 0.0001	0.134 (0.113)	0.540 (0.113)	t = 5.92 p < 0.0001	F = 0.13 ²¹ p = 0.719
1b) Compétence langagière générale à l'écrit: C-test écrit / mot correct	0.105 (0.081)	0.497 (0.081)	t = 4.22 p = 0.0003	0.077 (0.074)	0.448 (0.074)	t = 4.19 p = 0.0003	F = 0.03 p = 0.864
2) Correction grammaticale à l'écrit (potentiellement dérivée de règles): exercices pratiques de morphologie et de syntaxe	0.017 (0.181)	0.469 (0.181)	t = 3.79 p = 0.0013	0.386 (0.163)	0.829 (0.163)	t = 3.91 p = 0.0008	F = 0.003 p = 0.953
Partie orale I: Elicited imitation N = 131, n (IG) = 63, n (KG) = 68							
3a) Compréhension et restitution de langage oral: nombre d'éléments de contenu restitués	-0.143 (0.131)	0.236 (0.131)	t = 5.74 p < 0.0001	-0.159 (0.119)	0.301 (0.119)	t = 7.24 p < 0.0001	F = 0.776 p = 0.380 p _H = insignifiant
3b) Compréhension et restitution de langage oral: nombre de mots complets restitués	0.055 (0.110)	0.379 (0.110)	t = 5.89 p < 0.0001	0.025 (0.099)	0.495 (0.099)	t = 8.88 p < 0.0001	F = 3.64 p = 0.059 p _H = insignifiant
3c) Compréhension et restitution de langage oral: nombre de radicaux restitués	-0.126 (0.135)	0.229 (0.135)	t = 5.42 p < 0.0001	-0.164 (0.122)	0.276 (0.122)	t = 6.99 p < 0.0001	F = 0.89 p = 0.348 p _H = insignifiant
4) Caractère idiomatique de l'expression orale: <i>Chunking</i>	-0.099 (0.119)	0.098 (0.119)	t = 3.27 p = 0.007	-0.160 (0.108)	0.205 (0.108)	t = 6.31 p < 0.0001	F = 4.075 p = 0.046 p _H = insignifiant
5) Correction grammaticale (intuitive): correction de la morphosyntaxe	-0.029 (0.142)	0.248 (0.142)	t = 3.50 p = 0.0035	-0.196 (0.129)	0.455 (0.129)	t = 8.545 p < 0.0001	F = 11.566 p = 0.0009 p _H = 0.0045
Partie orale II: production orale N = 135, n (GE) = 64, n (GT) = 71							
6a) Fluidité: nombre de syllabes par durée d'enregistrement en secondes; vitesse d'élocution	1.490 (0.054)	1.610 (0.054)	t = 2.73 p = 0.035	1.527 (0.049)	1.516 (0.049)	t = -0.29 p = 0.99	F = 4.765 p = 0.031 p _H = 0.093
6b) Fluidité: nombre de syllabes par durée d'enregistrement en secondes, sans les réparations, répétitions et pauses remplies; vitesse d'élocution corrigée	1.190 (0.053)	1.331 (0.053)	t = 3.27 p = 0.0075	1.214 (0.048)	1.230 (0.048)	t = 0.39 p = 0.98	F = 4.42 p = 0.037 p _H = insignifiant
6c) Fluidité: nombre de syllabes par durée d'enregistrement sans les silences; vitesse d'articulation	3.006 (0.057)	3.075 (0.057)	t = 1.35 p = 0.53	3.118 (0.051)	3.086 (0.051)	t = -0.65 p = 0.92	F = 2.02 p = 0.157 p _H = insignifiant

Tableau 3 : Vue d'ensemble des résultats des tests

Si l'on tient compte de la correction de Holm (valeurs « p_H »), on n'observe de différence de progrès entre les deux groupes que dans un cas, à savoir le domaine 5: la correction de la morphosyntaxe dans la partie orale du test (*elicited imitation*). Pour ce domaine, les progrès du groupe témoin sont en moyenne plus grands que ceux du groupe expérimental. Si l'on ne tient pas compte de la comparaison multiple (valeur « p »), les différences suivantes sont significatives (en italique dans le tableau):

- à l'avantage du groupe témoin : 3b) nombre de mots complets restitués et 4) *chunking*
- à l'avantage du groupe expérimental : 6a) vitesse d'élocution (*speech rate*) et 6b) vitesse d'élocution corrigée (*pruned speech rate*)

Le groupe témoin affiche un progrès supérieur dans trois des cinq domaines évalués par l'*elicited imitation*, alors que dans les deux autres domaines, la progression des deux groupes est comparable. Il est intéressant de se pencher sur le contenu de ces différences: la progression du groupe témoin est plus forte dans les domaines où les critères « complet » et « correct » sont au premier plan (3b, 4 et surtout 5). Par contre, lorsque les aspects formels passent au second plan (3a, 3c), les différences entre les deux groupes deviennent négligeables.

Les progrès entre le test initial et le test final sont significatifs au sein de chacun des deux groupes pour presque toutes les mesures. Des exceptions existent toutefois: c'est le cas des mesures de fluidité, mises en évidence en gris clair dans le tableau. Pour le groupe témoin, aucune des mesures de fluidité n'a évolué entre t1 et t2, alors que pour le groupe expérimental, seule la vitesse d'articulation (*articulation rate*), soit la mesure de la pure et simple vitesse d'élocution à l'exclusion des pauses, est restée inchangée. Les moyennes de la vitesse d'élocution (*speech rate*) et de la vitesse d'élocution corrigée (*pruned speech rate*) ont par contre sensiblement augmenté de respectivement 8 et 12%.

19 Valeurs p pour test t du progrès par groupe selon la méthode de Tukey pour les comparaisons multiples.

20 Valeurs p pour test F avec correction pour *mixed-models* (Kenward-Roger-Approximation).

21 Pour des raisons de clarté, il n'y a aucune indication quant aux degrés de liberté. Ils peuvent être déduits à partir de la taille des groupes.

5.2 Résultats des entretiens avec les enseignants

Nous rapportons ici les principaux résultats des entretiens finaux avec les enseignants, classés selon les aspects cités au point 4.3.

Mise en œuvre du concept par les enseignants

Dans le but de garantir la comparabilité de l'enseignement entre toutes les classes du groupe expérimental et du groupe témoin, nous avons développé, introduit et fixé par écrit, sous forme d'un memento, un concept didactique pour chaque groupe. De plus, nous avons accompagné sa mise en œuvre par des conseils, des visites de classe et la consultation du journal des enseignants. Les entretiens permettent de conclure que l'enseignement a effectivement été fortement influencé par le concept et que les ressources didactiques propres ou les activités étrangères au concept n'ont joué qu'un rôle négligeable. L'élément novateur, consistant à débiter chaque journée de cours du groupe expérimental par 20 à 30 minutes d'entraînement à la fluidité, a été largement appliqué, même si cela s'est révélé difficile au début.

Acceptation des concepts par les participants et les classes

Le concept expérimental a reçu un accueil inégal dans les classes. Les formes plus dynamiques et sociales de certaines tâches orales de répétition se sont heurtées à la résistance de certains participants (plus âgés, par exemple), alors que d'autres apprenants les ont appréciées. L'une des enseignantes attribue l'ambiance particulièrement constructive de sa classe aux interactions sociales.

Le concept expérimental semble avoir été accepté plus facilement par les apprenants faibles et peu scolarisés que par les participants ayant un bon bagage scolaire ou voulant améliorer la correction de la langue. Pour les premiers, la suppression de l'enseignement grammatical a constitué un soulagement. Les seconds, eux, ont en partie douté de l'utilité d'un entraînement à la fluidité où les erreurs n'étaient pas ou peu corrigées. Ils ont également exprimé le souhait de recevoir un enseignement formel de la grammaire ou des ressources didactiques adaptées. Ils ont cependant reconnu que la répétition régulière d'éléments connus lors de l'entraînement à la fluidité était une occasion d'apprendre.

Dans les classes du groupe témoin, les avis négatifs ont surtout porté sur l'apprentissage du vocabulaire par listes de mots, tel que prévu par les manuels.

Dans les deux groupes, les dialogues de la vie quotidienne, enregistrés sur CD spécialement pour cette étude, ont rencontré un écho très favorable. Ces enregistrements ont été conçus comme une ressource supplémentaire, selon une approche didactique propre à chaque concept.

Ressenti des enseignants : avantages et inconvénients des concepts

Plusieurs enseignants des deux groupes ont trouvé leur marge de manœuvre très limitée par l'obligation de suivre un concept et des ressources déterminées. Cela a surtout été le cas des enseignants ayant travaillé avec le manuel moins spécifiquement adapté au groupe cible : ces derniers auraient souhaité pouvoir parfois laisser de côté certains éléments didactiques (thématiques, vocabulaire, progression grammaticale) et les remplacer par leurs propres ressources. En revanche, les enseignants du groupe expérimental ont clairement indiqué que la mise en œuvre du concept n'aurait pas été possible sans les ressources didactiques adaptées.

Lors des entretiens, les enseignants du groupe expérimental se sont montrés tantôt positifs, tantôt critiques quant aux diverses composantes du concept expérimental. L'importance accordée à l'oralité, surtout au travers de l'entraînement à la fluidité, a été saluée. Par contre, les enseignants ne sont pas certains que la langue écrite soit suffisamment promue. Ils estiment également que l'aspect individuel de l'entraînement à la fluidité est problématique, car les participants font et répètent fréquemment des erreurs et les enseignants n'ont pas suffisamment l'occasion de corriger.

Les activités centrées sur les *chunks*, qui remplacent dans le concept expérimental les tâches axées sur la grammaire, sont généralement appréciées. En même temps, le manque d'un enseignement explicite de la grammaire est considéré comme problématique : en effet, les apprenants au bénéfice d'un bon bagage scolaire ne reçoivent ni l'enseignement grammatical ni les ressources qu'ils attendent. Quant aux enseignants, ils déplorent la perte du métalangage grammatical qui constitue un outil de correction et d'explication efficace.

Le carnet de notes lexical individuel (*lexical notebook*) est globalement apprécié. Etant prévu comme base au travail individuel de répétition, il exige toutefois un travail de correction supplémentaire, puisqu'il faut corriger les entrées que les apprenants rédigent eux-mêmes. Pour les apprenants peu scolarisés d'ailleurs, le soutien requis de la part des enseignants est (trop) important. En revanche, les éléments individualisés comme le carnet ou les fiches de dialogues de la vie quotidienne sont vus de façon positive parce qu'ils donnent une plus grande autonomie d'apprentissage aux participants.

Les enseignants estiment que, contrairement au concept expérimental, le concept témoin (soit l'enseignement tel que le prévoit le manuel d'origine) s'adresse plus particulièrement à des apprenants rapides et au bénéfice d'un bon bagage scolaire. L'appréciation qu'ils en font est plutôt nuancée. D'un côté, les enseignants parlent d'une concentration excessive sur la grammaire et les concepts grammaticaux, ainsi que de trop de travail explicite sur le vocabulaire. D'un autre côté, ils reconnaissent l'importance d'un

travail approfondi sur le lexique. L'une des formatrices est d'avis que le concept témoin devrait être maintenu afin d'habituer les élèves peu scolarisés à l'apprentissage.

6 Discussion

Les résultats de l'expérience constituent une mine d'informations sur l'application pratique des deux concepts et sur les progrès linguistiques des participants.

Les entretiens avec les enseignants offrent des points de repères pour les trois premières questions de recherche (cf. chap. 3). Le nouveau concept didactique et les ressources nécessaires à sa mise en œuvre (question 1) semblent avoir été développés, appliqués et introduits avec succès. Lors des entretiens, certains enseignants des deux groupes ont cependant regretté le recours à une méthode prescriptive pour les cours lors de la phase expérimentale. Ils auraient souhaité recourir plus souvent à des ressources propres, plus directement adaptées aux besoins des participants. Du point de vue de la recherche cependant, cette méthode était nécessaire dans un groupe comme dans l'autre : sans elle, il aurait été plus difficile d'attribuer avec certitude les résultats de l'étude à l'influence de chacun des deux concepts.

L'acceptation du concept expérimental par les personnes concernées (question de recherche 2) est variable et l'évaluation doit être nuancée. En excluant des éléments normalement présents dans ce type de cours (not. le traitement explicite de la grammaire) et en mettant l'accent sur d'autres éléments (not. l'entraînement quotidien à la fluidité et des exercices sur les *chunks*, placés dans un cadre fonctionnel, mais sans analyse explicite), le concept va à l'encontre des attentes des apprenants ayant déjà participé à des cours de langue étrangère traditionnels et de celles des apprenants qui participent au cours dans le but d'améliorer la correction de leur expression. Par ailleurs, le concept expérimental met au défi la pratique préalable des enseignants. L'entraînement à la fluidité semble avoir posé problème aux enseignants ainsi qu'à une partie des participants parce qu'on y automatisait du langage pas encore suffisamment maîtrisé. L'idée de l'entraînement à la fluidité est d'utiliser uniquement du langage déjà connu, ce qui revient essentiellement à dire qu'il est maîtrisé (Nation & Meara, 2002). Toutefois, les participants semblaient ne réellement maîtriser que trop peu de langage pour les performances exigées et produisaient de nombreuses erreurs. L'absence d'enseignement explicite de la grammaire est également jugée problématique. Les participants ayant un bon bagage scolaire demandaient parfois que la grammaire soit enseignée de façon plus explicite et/ou souhaitaient des ressources didactiques adaptées, alors qu'il manquait aux enseignants la terminologie grammaticale pour faire des *feedbacks* et des corrections. En revanche, les apprenants peu scolarisés et plus faibles semblent avoir apprécié l'absence d'enseignement de la grammaire, ce qui réduisait également leur charge de travail. D'ailleurs, la réponse à la troisième question de recherche est plutôt positive : à en juger par les évaluations des enseignants, le concept expérimental est effectivement mieux adapté aux besoins des participants peu scola-

risés. Une des difficultés didactiques des cours à bas seuil, donc accessibles aux débutants, est généralement l'hétérogénéité des participants. Dans le cadre de cette étude, cette hétérogénéité s'est particulièrement révélée dans les divergences d'opinions face à l'absence d'enseignement explicite de la grammaire. Par conséquent, il pourrait être envisagé d'avoir spécifiquement recours au concept expérimental pour des apprenants peu scolarisés. Une autre option serait de compléter le concept expérimental par certains éléments plus conventionnels, en particulier le travail sur la grammaire, qui offre notamment une terminologie adaptée aux corrections et la possibilité d'approfondir le travail individuel au moyen de ressources didactiques adéquates. A l'inverse, l'entraînement à la fluidité, proposé par Nation & Meara (2002), pourrait être intégré au concept traditionnel : d'une part, parce qu'en mettant l'accent sur la correction, le risque d'automatiser des erreurs serait réduit et d'autre part, parce que l'acquisition ciblée de la fluidité est pratiquement exclue des programmes conventionnels, alors même qu'elle est reconnue comme un domaine de développement important et un aspect central de la compétence langagière.

Les résultats des tests de langue correspondent en partie aux attentes (optimistes) que nous avons formulées sur la base de la littérature spécialisée quant aux effets d'une approche basée sur les *chunks* (cf. chap. 3).

Nos attentes ont été confirmées dans trois domaines :

- La fluidité a progressé de façon plus importante dans le groupe expérimental.
- Comme prévu, il n'y a pas de différence en matière d'acquisition de vocabulaire (point de vue sémantique) et de capacité à communiquer des contenus.
- La correction de l'expression écrite progresse de manière égale dans les deux groupes.

Nos attentes n'ont pas été confirmées dans deux autres domaines :

- Le groupe témoin a fait plus de progrès dans l'utilisation de *chunks* aussi corrects et complets que possible que le groupe expérimental.
- La correction grammaticale à l'oral a plus progressé dans le groupe témoin.

Il est intéressant d'interpréter les résultats dans leur ensemble et à la lumière des entretiens avec les enseignants.

La question se pose de savoir dans quelle mesure le résultat global des tests correspond à l'observation des enseignants sur l'entraînement à la fluidité ? Selon eux, lors de cette activité récurrente au caractère essentiellement répétitif, le langage utilisé comportait souvent des erreurs. La répétition d'un langage incorrect amène-t-elle à mémoriser et produire des *chunks* ou constructions incorrects ? En tout cas, les résultats des tests ne contredisent pas cette possibilité : l'expression orale du groupe expérimental comporte en moyenne plus de fautes que celle du groupe témoin. C'est ce que montre la comparaison des trois domaines de l'*elicited imitation* où la correction joue un rôle : la correction grammaticale, la précision lexicale et l'utilisation de *chunks* corrects. En parallèle, la fluidité à l'oral s'est améliorée dans le groupe expérimental plus que dans le groupe témoin. Cela se traduit non pas par une accélération de l'articulation en soi (*articulation rate*), mais par une réduction des pauses (soit pauses, remplissages, répétitions et réparations) sur le temps d'expression total, comme l'indiquent les résultats des mesures de vitesse d'élocution (*speech rate*) et de vitesse d'élocution corrigée (*pruned speech rate*). Dans le groupe témoin, les pauses sont plus longues et le temps d'expression donc moins productif, ce qui indique un besoin de planification ou de correction plus important. Par contre, les participants du groupe expérimental semblent pouvoir recourir à plus de modèles existants, signe que la langue apprise appartient déjà au savoir procédural (de Jong & Perfetti, 2011) ou qu'ils disposent de constructions. Les mesures de la correction écrite, qui ne laissent transparaître aucune différence entre les deux groupes, ne contredisent pas cette interprétation. En effet, il est possible que sous la relative pression temporelle de l'*elicited imitation*, le caractère inarrêtable (balistique) du langage (Segalowitz & Hulstijn, 2009) se soit manifesté, alors que l'examen écrit laissait le temps de s'autocorriger en recourant à d'autres connaissances. Reste à déterminer de quelles formes de savoir il s'agit là : les participants ont-ils comparé plusieurs *chunks* assimilés ou ont-ils fait appel à des connaissances abstraites (règles) qu'ils auraient déduites de leur expérience ?

Les résultats des tests oraux devraient être examinés sous l'angle des résultats de recherche sur la performance relative aux tâches d'expression, notamment en tenant compte de l'hypothèse du *tradeoff* (Skehan, 2009). Cette hypothèse part du constat largement reconnu que nos capacités cognitives sont limitées. Ainsi, notre mémoire de travail et notre capacité d'attention sont restreintes. Lorsque l'on parle naît donc une situation de concurrence entre les exigences de complexité, de correction et de fluidité.²²

Les tâches de nos tests étant très structurées, le degré de complexité n'était probablement pas un facteur important et il était en principe le même pour les deux groupes.

Les ressources cognitives pouvaient par conséquent être déployées en priorité pour la correction et la fluidité. Il apparaît que le groupe expérimental a donné la préférence à la fluidité, alors que le groupe témoin a privilégié la correction. Selon les partisans de l'approche lexicale (Nation & Newton, 2009), cela ne devrait pas être le cas. On s'attendrait plutôt à ce que les *chunks*, éléments complets à disposition, permettent d'augmenter non seulement la fluidité mais encore la correction. Dans le cas où les *chunks* déjà acquis requerraient encore une adaptation grammaticale, on pourrait s'attendre à ce que le recours à des *chunks* facilement et rapidement mobilisables libère des ressources cognitives pouvant alors être entièrement consacrées à l'adaptation ou la correction nécessaire. Pourtant, rien de cela n'apparaît dans les données. Il est concevable qu'ici, les résultats trahissent non pas les limites des ressources cognitives pendant la production langagière, mais bien plus les différentes caractéristiques des savoirs et compétences développés dans les deux groupes pendant les cours. Le groupe expérimental dispose probablement de *chunks* certes incorrects, mais plus automatisés, produits par la mémoire sous l'effet du caractère inarrêtable (balistique) de la langue, alors que le groupe témoin dispose des connaissances linguistiques et de l'attention nécessaires pour produire un allemand oral un peu plus correct, au prix toutefois de la fluidité.

Nos résultats vont aussi à l'encontre de la position de *non-interface* (Krashen, 1982) : l'acquisition d'une grammaire libre et spontanée semble justement avoir mieux réussi grâce à l'apprentissage explicite. Cependant, le cours du groupe expérimental, dans lequel les phénomènes grammaticaux à acquérir apparaissent systématiquement plusieurs fois, a entraîné une amélioration comparable dans l'utilisation écrite de la langue. Il s'est visiblement produit un apprentissage fortuit grâce auquel les apprenants ont pu disposer de ressources grammaticales dans lesquelles puiser pour effectuer des tâches écrites.

Mais que dire de la supposition selon laquelle le groupe ayant reçu en enseignement basé sur les *chunks* s'appuierait sur ces derniers pour faire l'analyse grammaticale de la langue déjà utilisée ? Ce scénario est en principe envisageable. Il dépend de ce que signifie, sur le plan cognitif, la différence entre la correction orale et la correction écrite observée dans le groupe expérimental. Les progrès relativement importants à l'écrit signifient-ils que des formes correctes ont effectivement été mémorisées, mais sont moins automatisées et ne revêtent pas encore ce caractère inarrêtable (balistique)

22 Les liens entre complexité, correction et fluidité (*complexity*, *accuracy* et *fluency*, abrégé CAF) font actuellement l'objet d'un débat en linguistique appliquée et dans la recherche sur l'évaluation des compétences linguistiques.

de la langue ? Dans ce cas, il est fort possible qu'une analyse des formes linguistiques puisse être faite dans le cadre d'un autre cours, dont objectif serait l'enseignement d'un langage formellement correct.

Une observation importante a été faite grâce aux deux mesures de vitesse d'élocution : la fluidité n'a progressé que chez les apprenants du groupe expérimental. Elle confirme une supposition de Lewis (1993) et d'autres défenseurs de l'approche lexicale. La fluidité est l'un des domaines de compétences fondamentaux à développer. Or, le groupe témoin semble avoir stagné sur ce point. Reste à savoir si cela n'est qu'un symptôme du phénomène de concurrence entre correction et fluidité, découlant des limites de capacité de la mémoire. Cette supposition est plausible dans la mesure où les résultats de mesure de la vitesse d'élocution corrigée (*pruned speech rate*), qui reflètent également des aspects de planification et d'autocorrection, indiquent que les activités d'expression demandaient plus de travail au groupe témoin qu'au groupe expérimental. À l'avenir, un des objectifs d'apprentissage de ce groupe devrait être de davantage développer le langage automatisé (*chunks* et constructions) par le biais d'activités de répétition. Une chose est certaine : les contacts sociaux sont plus faciles lorsque l'expression est fluide. À long terme, la fluidité améliore donc les chances de continuer à apprendre en dehors du cours.

7 Bilan

Au vu des résultats, il convient de répondre de façon nuancée à la question pratique qui se trouve à l'origine de cette étude. Un enseignement axé sur les *chunks*, renonçant à un travail explicite sur la grammaire (*focus on forms* et *focus on form*,²³ Long, 1991) peut certes être mis en œuvre avec succès dans des cours de langue à bas seuil, mais il se heurte à certaines réserves chez les apprenants ayant déjà suivi des cours de langue étrangère et chez ceux qui souhaitent améliorer la correction de la langue. En revanche, selon les observations des enseignants, ce concept semble faciliter la tâche des apprenants peu scolarisés. Les enseignants, quant à eux, estiment que la réussite d'un cours basé sur le concept expérimental nécessite des ressources adaptées. Ils soulignent aussi l'aspect problématique de la répétition régulière, apparemment inévitable, d'un langage peu correct et authentique. Ce point pourrait expliquer le fait que le groupe expérimental a fait moins de progrès en matière de correction orale que le groupe témoin, lequel a suivi un enseignement procédant d'un concept traditionnel qui fait intervenir de nombreuses activités axées sur les formes. À l'inverse, dans le domaine tout aussi important de la fluidité orale, c'est le groupe témoin qui n'a pas fait de progrès mesurables pendant la phase d'expérimentation. Ces observations parlent en faveur de l'implémentation de *balanced learning opportunities* (Ellis, 2005b). Pour le concept expérimental, il s'agirait de davantage se concentrer sur la correction des *chunks* devant être automatisés. Cela ne signifie pas que l'on doive nécessairement faire un travail systématique sur la grammaire, mais que l'accent soit placé de façon générale sur les formes linguistiques, c'est-à-dire que l'on insiste sur les formes correctes dans la langue cible, qui seront ensuite répétées dans le cadre de l'entraînement à la fluidité. Pour le concept témoin, par contre, il s'agirait d'encourager l'automatisation de la langue au travers d'activités de répétition.

D'un point de vue scientifique, il serait souhaitable d'analyser en profondeur les productions orales qui ont permis d'évaluer la fluidité, pour en apprendre davantage sur le travail de formulation et ses différences d'un groupe à l'autre, et donc sur la base cognitive des productions. Existe-t-il des différences spécifiques dans le type d'erreurs formelles ? Le type de pauses remplies, de réparations et de répétitions présente-t-il des différences ? Il serait en outre intéressant d'examiner l'effet d'un entraînement ciblé à la fluidité dans un cadre scolaire classique, où l'enseignement n'est pas intensif et où les contacts linguistiques hors cadre scolaire sont rares. Ce genre d'étude devrait couvrir

23 *Focus on forms* : enseignement explicite systématique de la grammaire ; *focus on form* : enseignement explicite de la grammaire sur la base des explications requises par un exercice orienté sur la communication.

plusieurs niveaux de compétences, pour tenir compte du fait que la taille du répertoire de formes correctes que détient un individu avant l'entraînement à la fluidité pourrait expliquer les différences de résultats dans le développement de l'exactitude formelle.

Fluenza e correttezza – fluenza o correttezza?

Risultati di uno studio d'intervento nei corsi di base intensivi di tedesco come lingua seconda per principianti

—
Peter Lenz, Malgorzata Barras, Fabienne Manz¹

¹ Un particolare ringraziamento a Thomas Studer, responsabile della ricerca, per le sue preziose proposte sull'esecuzione del progetto e sulla presentazione di questo rapporto.

Riassunto

Il presente studio si interroga sull'efficacia di un insegnamento basato principalmente sull'approccio lessicale (*lexical approach*) (Lewis, 1993; 1997), focalizzato sull'apprendimento di *chunk* (sequenze di parole con parti variabili) e sulle esercitazioni di fluenza (Nation & Meara, 2002) in classi di tedesco L2 con apprendenti principianti adulti in parte poco abituati all'apprendimento formale.

All'interno di un setting simil-sperimentale, costituito da un gruppo d'intervento e un gruppo di controllo (GI e GC), dodici classi suddivise in sei coppie hanno appreso il tedesco L2 seguendo un manuale. Nelle sei classi composte dal GI, la struttura e i materiali didattici delle lezioni seguono un paradigma didattico orientato allo studio d'intervento appena citato, mentre nelle sei classi composte dal GC, l'insegnamento si attiene principalmente alla traccia prevista da un manuale, le cui caratteristiche principali sono un insegnamento esplicito della grammatica e un apprendimento del lessico con singole parole. Grazie a un test di entrata e un test finale, è stato possibile valutare differenti competenze orali e scritte. Alcune interviste con i docenti hanno inoltre permesso di chiarire questioni legate all'idoneità dei due piani di studio per i gruppi target ma anche di tematizzare particolarità nel processo d'insegnamento e apprendimento di entrambi i gruppi.

Dalle interviste emerge che, seguendo entrambi i piani di studio, è stato possibile insegnare con profitto. Gli apprendenti del GI con esperienze pregresse nell'apprendimento di LS e tendenti a volersi esprimere correttamente hanno tuttavia lamentato in parte la mancanza dell'insegnamento della grammatica e di materiali didattici su temi grammaticali. Essi hanno tuttavia riconosciuto nelle esercitazioni di fluenza delle buone occasioni d'apprendimento. Le esperienze dei docenti confermano che il piano di studi del GI è più adatto a apprendenti non abituati a studiare una lingua, mentre quello del GC si presta piuttosto a chi è abituato all'apprendimento formale. I docenti del GI hanno sentito la mancanza della metalingua grammaticale in occasione della correzione di errori; inoltre, hanno avuto poche occasioni per correggere errori durante le esercitazioni di fluenza: gli errori erano così ripetuti e parzialmente interiorizzati.

I risultati dei due test che valutano lo sviluppo delle diverse competenze linguistiche mostrano, per quasi tutti gli apprendenti, un'evoluzione positiva. Singole eccezioni si registrano unicamente nelle misure di fluenza: nel GC, diversamente che nel GI, i due indicatori di fluenza principali non mostrano, infatti, un incremento. I progressi della *speech rate* (numero di sillabe al secondo nell'arco di tempo complessivo di elocuzione incluse tutte le pause) e la *pruned speech rate* (corrisponde alla *speech rate*, ma esclude le sillabe delle pause piene, ripetizioni e autocorrezioni) sono forti indicatori del fatto che gli apprendenti del GI hanno svolto compiti di produzione orale con

minore dispendio di risorse sul fronte della pianificazione e dell'autocontrollo rispetto agli apprendenti del GC. I risultati dettagliati delle parti dell'esame orale portano inoltre a pensare che un incremento di fluenza implichi una perdita di correttezza linguistica. Sul piano funzionale e di contenuto, invece, le competenze di produzione orale dei due gruppi non si differenziano. Risultati analoghi si osservano nell'ambito della produzione scritta, anche a livello della morfosintassi, dove si prevedevano differenze a vantaggio del GC. Nel presente rapporto, infine, i risultati differenziati dell'analisi tra i due piani di studio vengono non solo discussi alla luce dei paradigmi teorici relativi all'apprendimento e alla produzione linguistica, ma anche sfruttati per trarne delle conseguenze per ulteriori ricerche.

1 Introduzione

Il progetto “Impatto dei metodi didattici sull’apprendimento dei migranti adulti nei corsi di lingua seconda”, all’interno del quale il presente studio si posiziona, è stato finanziato nell’ambito del programma di ricerca 2011-2015 del Centro scientifico di competenza per il plurilinguismo e dalla Segreteria di Stato della migrazione (SEM). Il progetto mirava a far luce, grazie a dati empirici, su una questione discussa da diversi esperti nel processo di elaborazione del “Curriculum di riferimento per la promozione delle conoscenze linguistiche dei migranti” svizzero (Lenz, Andrey & Lindt-Bangerter, 2009). Si trattava di sapere se, nel contesto di un insegnamento di tedesco L2 a studenti parzialmente senza esperienza di apprendimento formali, convenisse ricorrere al paradigma didattico tradizionale (con un insegnamento della grammatica esplicito e analitico, ovvero un *focus on forms* secondo Long (1991) e un lavoro sul lessico focalizzato sull’apprendimento di singole parole), oppure se fosse da preferire un paradigma volto all’apprendimento e all’automatizzazione di sequenze linguistiche non analizzate (*chunk*). Il presente studio tenta di rispondervi tramite osservazioni basate su prove concrete tratte dall’uso in classe.

A tale scopo è stato scelto un design simil-sperimentale con un gruppo d’intervento e un gruppo di controllo. Sei coppie di classi sono state così sottoposte a un insegnamento orientato al principio dei *chunk*, mentre sei altre classi a un insegnamento fortemente orientato all’apprendimento della grammatica e di parole singole, come previsto dal manuale in uso.

Inizialmente, hanno partecipato alla sperimentazione quattro scuole di lingua offrendo corsi di tedesco L2 di base. Lo studio è stato concluso nei termini previsti in tre scuole per un totale di 12 classi. La quarta scuola ha abbandonato il progetto dopo la fase iniziale perché la grande eterogeneità delle proprie classi ha richiesto uno smistamento dei partecipanti. La collaborazione con le scuole, grazie al grande impegno di tutte le persone coinvolte, è stata ottima.²

Un elemento fondamentale per l’implementazione dello studio è rappresentato dai materiali d’insegnamento: sono stati scelti manuali di tedesco L2 regolarmente usati dalle scuole stesse. Due scuole usano lo stesso manuale, una terza un altro.³ Per adat-

2 Un cordiale ringraziamento va alle seguenti scuole: aoz (Zurigo), FORMAZIONE (Berna), Lernpunkt (Berna) e ECAP (Lucerna).

3 Si tratta delle edizioni svizzere dei manuali *Schritte plus 1. Niveau A1/1. Ausgabe Schweiz* (Niebisch, Penning-Hiemstra, Specht & Bovermann, 2011) della casa editrice Hueber e di *Studio d A1. Schweizer Ausgabe* (Funk, Kuhn, Demme & Bayerlein, 2005) della casa editrice Cornelsen.

tare i materiali didattici di tali pubblicazioni al paradigma didattico orientato all’apprendimento in *chunk*, è stato necessario adattarne e integrarne diverse parti per poi stamparne una dispensa rilegata.⁴

Studi come il seguente sono rari: sono poche le ricerche che si concentrano sull’apprendimento di migranti adulti, anche se la promozione delle competenze linguistiche nell’ambito della migrazione è un tema socialmente rilevante. Sono anche rari gli studi d’intervento basati su un approccio simil-sperimentale della durata di più settimane e con implicazioni dirette per i piani di studio dei corsi. Essi rappresentano una sfida non solo dal punto di vista pratico, ma anche dal punto di vista scientifico, perché diversi fattori che possono sminuire l’effetto dell’intervento sono meno controllabili che nel corso d’interventi puntuali. D’altro canto, tuttavia, sono proprio i dati elicitati nel quadro d’interventi integrati “nell’ecosistema insegnamento” che permettono di trarre delle conclusioni per eventuali applicazioni nella prassi didattica.

4 Ringraziamo coloro che ci hanno autorizzato ad adoperare il loro materiale didattico.

2 Situazione iniziale e basi teoriche

Il presente studio esamina contrastivamente lo sviluppo delle competenze linguistiche di adulti principianti in due diversi tipi di corsi intensivi di tedesco L2. L'obiettivo primario di entrambi i corsi risiede nell'incrementare l'abilità d'uso della lingua degli apprendenti, affinché essi possano partecipare con maggiore profitto alla vita sociale del paese d'accoglienza. Il nostro studio, tuttavia, non verifica l'abilità linguistica solo attraverso compiti comunicativi, ma si sofferma su una selezione di competenze linguistiche quali la fluenza e la correttezza grammaticale. L'accento da noi posto sulle risorse linguistiche è riconducibile alla tradizione di ricerca nell'ambito della linguistica applicata. Questo *modus faciendi* permette di ricavare dati maggiormente differenziati e più facilmente interpretabili e trova una sua giustificazione, tra l'altro, nel fatto che proprio per dei principianti, è di centrale importanza sviluppare un repertorio linguistico di base per l'uso attivo.

Raggiungere un buon grado di fluenza in una lingua è per molti l'obiettivo primario dell'apprendimento linguistico. Con fluenza si intende generalmente la capacità di utilizzare una lingua velocemente, facilmente e correttamente (Segalowitz, 2005). A livello cognitivo e fisiologico-articolatorio, la fluenza presuppone un determinato grado di automatizzazione. L'apprendimento di un comportamento complesso richiede attenzione, concentrazione e impegno. Normalmente le prestazioni sono inizialmente lente e erranee, ma migliorano attraverso la ripetizione fino a essere prodotte con facilità (Aguado, 2002). L'automatizzazione non è unicamente riconducibile a una velocizzazione dei processi linguistici, ma implica anche una ristrutturazione dei processi nella memoria (Anderson et al., 2004; Gatbonton & Segalowitz, 2005), che, a loro volta, riducono il carico per la memoria attiva durante la produzione linguistica. Sono segnali di automatizzazione una maggiore velocità, facilità, spontaneità e irrefrenabilità (irrefrenabilità del linguaggio, *ballistic*) nell'uso della lingua. Essi non devono per forza manifestarsi contemporaneamente (Segalowitz & Hulstijn, 2009). Secondo Segalowitz & Hulstijn, la *ballistic* è la prova inconfutabile del raggiungimento dell'automaticità linguistica. S'indica con il termine di *ballistic* il fenomeno per il quale l'uso automatico della lingua diviene relativamente immune da interferenze. Applicato all'acquisizione di una L2, il concetto di *ballistic* degli automatismi può essere considerato responsabile del fatto che il discente mantenga strutture linguistiche non conformi alle norme della lingua bersaglio a lungo o addirittura in modo permanente, ovvero che tali forme si fossilizzino⁵ (Bybee, 2008).

5 Long (2005) rimette in discussione la "fossilizzazione" (permanente) come fenomeno cognitivo reale. Nella sua analisi critica di studi, mostra che questo temuto stato di

Se l'automatizzazione e quindi l'automaticità sono cruciali per lo sviluppo della lingua, sorge la domanda di come conseguirle nei corsi di lingua. Dal punto di vista della linguistica *usage-based* (basata sull'uso), si tratta di instaurare e utilizzare processi naturali che conducano a un uso automatizzato della lingua. Lo scopo dell'insegnamento e dell'apprendimento consiste nel far nascere e sviluppare nei discenti cosiddette *construction*. Il presupposto di base è che chi apprende una lingua (L1 o lingue successive) evince elementi e convenzioni dell'uso della lingua a partire dalla propria esperienza di comunicazione. Queste nozioni e competenze esperienziali vengono memorizzate come *construction* in seguito a un processo di astrazione (proceduralizzazione). Grazie alle *construction* i discenti possono far corrispondere alcune funzioni (comunicative) a determinate forme linguistiche (Ellis, 2005a). Composte da *chunk* (cognitivi) fissi (*neuromotor routines*) e da una quantità variabile di elementi intercambiabili (Bybee, 2008), le *construction* possono essere parole, espressioni o intere frasi. Un esempio di *construction* con uno spazio vuoto che può accogliere elementi intercambiabili è l'espressione semi-fissa "gradirei...". Prima di poter adoperare una *construction* in un enunciato, questa dev'essere riconosciuta in quanto tale e automatizzata attraverso la ripetizione. Il locutore potrà esprimersi fluentemente non appena sono disponibili *construction* automatizzate in quantità sufficiente, soprattutto *construction* estese che ne contengono a loro volta di più piccole. Le *construction* consentono di sgravare la memoria attiva con le sue capacità limitate e sbloccano risorse cognitive impiegabili, per esempio, per gestire attivamente la lingua (Skehan, 2009).

Bybee (2008) propone riflessioni riguardo a un curriculum in grado di promuovere lo sviluppo di *construction*: è fondamentale creare condizioni di acquisizione che compensino il minore contatto che il discente ha con la L2 rispetto alla L1. Per cominciare, bisogna operare una selezione mirata di sequenze linguistiche sulla base dell'analisi di corpora. Dalle sequenze, per mezzo di esempi prototipici della lingua, si dovrebbero poter ricavare importanti *construction*. Allo scopo di rendere i modelli riconoscibili in quanto tali e avviare l'acquisizione, i discenti devono essere confrontati ripetutamente agli stessi modelli e in modo mirato a modelli modificati.

Secondo Segalowitz & Hulstijn (2009), un insegnamento di questo genere, benché raccomandabile da una prospettiva linguistica per lo sviluppo dell'automaticità, nella

difetto definitivo nell'acquisizione di una lingua (straniera) spesso non sussiste là dove viene diagnosticato. Le caratteristiche linguistiche non conformi, indicano più che altro un'interruzione del processo di acquisizione della lingua, il che può essere dovuto a varie ragioni.

pratica rischia di essere osteggiato. Sebbene la ripetizione aiuti a generare automaticità, i discenti potrebbero considerarlo un esercizio troppo decontestualizzato e poco interessante. I tipici corsi comunicativi appaiono invece perlopiù come pertinenti dal punto di vista contenutistico, ma prevedono una quantità di ripetizioni insufficiente a provocare l'automaticità. Gli autori propongono quindi di privilegiare concetti orientati alle esigenze di comunicazione in cui si imparino enunciati della lingua bersaglio molto versatili, che saranno poi esercitati a sufficienza nel quadro di attività comunicative e rilevanti dal punto di vista del contenuto.

Sul piano didattico, i nostri concetti per il gruppo d'intervento e per il gruppo di controllo si differenziano principalmente per l'approccio grammaticale. Nel gruppo di controllo, la grammatica viene insegnata in maniera esplicita e sistematica (*focus on forms*), mentre nel gruppo d'intervento non viene trattata espressamente come argomento del corso. Si pone quindi una questione da tempo dibattuta nella ricerca sull'acquisizione linguistica e nella glottodidattica: le nozioni linguistiche che vengono veicolate in modo esplicito, mediante regole, e poi esercitate sono attivabili in modo altrettanto spontaneo dai discenti quanto le nozioni che hanno dedotto da soli a partire dall'input della lingua straniera? Su questo argomento, è nota l'opinione radicale di Krashen (1982), ossia l'ipotesi della non interazione (*non-interface position*): le nozioni fondate su regole verrebbero attivate soltanto con una finalità di sorveglianza dell'uso della lingua. Oggi, si è affermata l'opinione che processi di apprendimento complessi come l'acquisizione di una lingua possono senz'altro essere impostati in modo cosciente (come sapere dichiarativo) per poi evolvere in comportamenti automatizzati (in quanto elementi del sapere procedurale; cfr. il noto modello generale dello sviluppo cognitivo di Anderson et al., 2004). Da un ampio meta-studio sul successo di diverse forme di corso (Norris & Ortega, 2001) emerge d'altronde che un insegnamento in cui viene prestata un'attenzione – sistematica o puntuale – a fenomeni linguistici dà esiti migliori rispetto a un insegnamento puramente implicito. Un insegnamento che include anche un sapere linguistico esplicito sembra produrre progressi particolarmente soddisfacenti. In un saggio più recente che considera diverse forme di verifica aperta e un maggior numero di risultati sull'insegnamento implicito, Spada & Tomita (2010) confermano la superiorità di tipologie d'insegnamento con una parte esplicita rispetto a quelle puramente implicite. Il nostro concetto d'intervento comprende varie forme di esercizio focalizzate su determinati fenomeni linguistici ma questi fenomeni non vengono affrontati in modo esplicito, per cui la differenza rispetto al concetto del gruppo di controllo – decisamente più improntato sull'analisi grammaticale esplicita – è evidente.

Il nostro concetto d'intervento è stato prevalentemente impostato sull'approccio di glottodidattica detto *lexical approach* di Michael Lewis,⁶ ben noto sin dagli anni Novanta del secolo scorso (p.es. Lewis, 1993; 1997; 2000) e che abbiamo completato in particolare con alcune proposte di Nation (2001) sull'esercitazione alla fluency.⁶

Lewis fonda il suo approccio soprattutto sui dati della linguistica dei corpora relativi alla frequenza di collegamenti più o meno stabili tra elementi lessicali. Secondo Lewis, essi costituiscono infatti la chiave della grammatica: "Language consists of grammaticalized lexis, not lexicalized grammar" (Lewis, 1993: 34). Consigliava quindi di insegnare partendo dai *chunk*, poiché la lingua è in gran parte formata da queste unità lessicali non manipolabili a piacimento, ossia non manipolabili secondo quanto sarebbe in sé permesso dalla grammatica. Introducendo il concetto dei *chunk*, Lewis (1997) individua due categorie principali: collocazioni ed espressioni. Le collocazioni sono parole che co-occorrono con grande frequenza (p.es. "scattare una fotografia") e non, come in altre lingue, "prendere una fotografia". Le espressioni vanno distinte, secondo Lewis, in fisse (*fixed*) (p.es. "buongiorno") e semi-fisse (*semi-fixed*) (p.es. "gradirei..."), queste ultime caratterizzate dal fatto di essere variabili in uno o più punti.

Nel nostro studio, definiamo come *chunk* sia le espressioni fisse sia quelle semi-fisse. Consideriamo espressioni fisse anche unità funzionali di una sola parola come "Grazie!". Nel concetto d'intervento, le collocazioni non formano una sottocategoria a sé stante, ma vengono introdotte come parti di espressioni.

Applicando l'approccio lessicale, sin dall'inizio si introducono nell'insegnamento *chunk* di diversa lunghezza, senza tuttavia analizzarli dal punto di vista grammaticale (Lewis, 1993). Essi vengono proposti tramite un insegnamento orientato alle capacità e alle esigenze reali dei partecipanti, affinché i discenti dispongano quanto prima di mezzi linguistici per comunicare nella vita quotidiana. Ci si aspetta che questo approccio stimoli la fluency, poiché mette a disposizione elementi linguistici più lunghi che i discenti possono richiamare per intero e in tal modo si favorisce la riuscita della comunicazione. Rispetto a un insegnamento più convenzionale, si bada poco all'apprendimento e all'esercitazione dell'analisi logica o alla memorizzazione arbitraria di singoli vocaboli. Vengono invece svolti molti esercizi di comprensione – dapprima soprattutto orale, poi scritta – e di apprendimento attivo mediante un glossario personale di *chunk* (*lexical notebook*) e quaderni didattici (Lewis, 1997).

L'approccio lessicale gode di un largo apprezzamento nella letteratura special-

6 Altri tentativi, meno noti, di applicare le conoscenze sui *chunk* ai corsi di lingua sono stati intrapresi da Willis (1990) e Nattinger & DeCarrico (1992).

stica. Nessuno nega che la lingua contenga un grande numero di *chunk*. Wray & Perkins (2000) fanno notare che la lingua di un adulto nativo (la cui L1 è l'inglese) è costituita per il 70% circa di espressioni formulari (incluse le collocazioni). Si è inoltre sovente fatta notare l'utilità dell'apprendimento di *chunk* per l'uso comunicativo della lingua sin dalle fasi iniziali di acquisizione. Se i *chunk* vengono appresi come lessico, in poco tempo possono essere impiegati come strumenti di comunicazione (corretta) a prescindere da lacune sul funzionamento della lingua da cui sono tratti. Se la comunicazione per mezzo di *chunk* riesce, aumentano le probabilità che si continui a imparare la lingua (Aguado, 2002; Nation & Newton, 2009; Wray, 2002). Si presuppone che attraverso l'apprendimento e l'uso precoci di *chunk* avvenga anche uno sviluppo della conoscenza implicita ed esplicita della struttura della lingua. Detto altrimenti, si presume che il repertorio di espressioni formulari apprese venga via via – o a partire da un determinato momento del processo – analizzato dal punto di vista delle regole grammaticali (p.es. Lewis, 1997; Wray, 2000). I risultati di questa analisi vengono poi memorizzati in aggiunta ai *chunk* e vi si potrà così attingere secondo le proprie esigenze di uso della lingua (Wood, 2002).

La questione centrale è dunque di verificare se tale analisi avvenga veramente e quando. Prendendo spunto dalla linguistica basata sull'uso, Bybee (2008) propone di completare l'approccio lessicale con elementi didattici che aiutino in modo mirato a sviluppare *construction* morfosintattiche.

Nel concetto di corso per il gruppo d'intervento, l'approccio lessicale è stato completato con un'esercitazione alla fluenza per favorire l'acquisizione linguistica mediante la ripetizione. Questa scelta si basa anche sul quadruplice concetto di Nation (*four strands approach*) per un buon apprendimento (p.es. Nation & Meara, 2002). Si tratta di dedicare il 25% del tempo dei corsi a attività linguistiche in cui parole già note vengono utilizzate senza prestare troppa attenzione alla lingua, ma imponendo in compenso una certa velocità di elocuzione, il che dovrebbe contribuire a rendere tali elementi disponibili al discente per un uso fluente. Si distinguono due varianti: nella prima, il materiale linguistico noto è ripetuto varie volte (ad esempio incrementando la pressione a livello di tempo); nella seconda, ci si concentra sull'arricchimento lessicale, producendo combinazioni e utilizzi nuovi dei vocaboli conosciuti (p.es. tramite lettura di testi più lunghi benché semplificati). Nel nostro concetto d'intervento, abbiamo attuato soltanto il primo aspetto.

3 Obiettivi dello studio e domande di ricerca

Un obiettivo dello studio è stato lo sviluppo di un concetto di corso valido per il gruppo d'intervento e la sua attuazione con del materiale didattico adattato. Si sono inoltre dovuti trovare organizzatori di corsi di lingua disposti a lavorare tra i due e i quattro mesi secondo le nostre indicazioni. L'aspetto che più interessava in questo studio era il progresso dei partecipanti ai corsi, che andava rilevato e analizzato con strumenti adeguati.

La nostra ricerca si prefigge di rispondere alle seguenti domande:

- i. è possibile preparare un concetto d'insegnamento efficace e il materiale occorrente per l'intervento e poi adoperarli nel quadro di un corso di base di tedesco come lingua straniera?
- ii. gli insegnanti e i partecipanti ai corsi accettano il concetto d'intervento e lo trovano interessante?
- iii. il concetto d'intervento è adatto ai partecipanti, soprattutto a quelli con poca familiarità con l'apprendimento scolastico?
- iv. in che misura i progressi dei partecipanti del GI differiscono da quelli dei partecipanti del GC nei diversi ambiti?

Le risposte alle prime tre domande sono state ottenute intervistando gli insegnanti.

I progressi dei partecipanti sono stati misurati con test orali e scritti. Il confronto dei due gruppi è stato effettuato secondo l'ipotesi che il concetto d'intervento è da considerare più efficace quando, paragonato al GC, il GI ha raggiunto i seguenti risultati:

- iv.a. un impiego più frequente di *chunk* della lingua bersaglio
- iv.b. una maggiore fluenza linguistica
- iv.c. una correttezza relativamente più elevata nell'uso orale (correttezza intuitiva) e uguale o solo di poco inferiore nell'uso scritto (correttezza nell'applicazione di regole)
- iv.d. una ricchezza di lessico paragonabile (a livello di significati) e una capacità analoga di comunicare contenuti (a prescindere dalla correttezza formale)

4 Metodo

4.1 Design della ricerca

Per questo studio è stata scelta un'impostazione simil-sperimentale con un gruppo d'intervento e un gruppo di controllo. Abbiamo lavorato unicamente sui corsi di base⁷ intensivi, quindi corsi in cui gli studenti partecipavano a lezioni di tedesco durante 4-5 giorni a settimana per un totale di almeno otto e al massimo undici ore settimanali. A seconda degli organizzatori, erano previste tra le otto e le sedici settimane per completare il corso con il manuale.

Allo scopo di compensare gli effetti di questa e di altre differenze tra i vari organizzatori, prima di cominciare, i principianti⁸ iscritti sono stati divisi in coppie di classi, che avrebbero seguito le lezioni secondo l'uno o l'altro concetto. L'assegnazione al gruppo d'intervento o a quello di controllo è avvenuta sulla base di un campionamento casuale stratificato. L'esperienza specifica degli organizzatori con questo gruppo target è stata sfruttata per tentare di mantenere costanti in ognuna delle due classi le caratteristiche potenzialmente rilevanti per l'apprendimento, ossia in particolar modo l'origine, la lingua materna e la familiarità con l'ambiente scolastico.

Le classi così formate sono state assegnate dagli organizzatori dei corsi agli insegnanti. In tal modo, si sono potute formare sette coppie di classi presso quattro organizzatori diversi. Una coppia si è dovuta spaiare, perché la composizione delle classi era troppo eterogenea a causa dei nostri principi di assegnazione, creando insoddisfazione sfociata nella necessità di suddividere gli studenti in gruppi di diverso livello.

Il nostro intervento si è svolto in due parti di uguale lunghezza. Nella prima, quella preparatoria, molti partecipanti hanno acquisito le primissime nozioni di tedesco e gli insegnanti si sono abituati a lavorare con i nostri concetti e materiali. Dopo questa fase, ai fini dello studio, si è svolto un test d'entrata, la cui struttura è in seguito stata ripresa quasi identica nel test finale. Alla fine della prima metà del corso, tra i partecipanti ci sono state alcune partenze e alcuni nuovi arrivi. I nuovi partecipanti sono stati sottoposti allo stesso test d'entrata. Il test finale si è di regola tenuto nell'ultima settimana della seconda metà del corso. L'analisi quantitativa concerne quindi l'evoluzione della seconda metà del corso.

7 L'accesso ai corsi di base, è facilitato su più aspetti (livello previo di conoscenze scolastiche e linguistiche, finanziamento).

8 Questa designazione si fonda sui livelli ai quali gli interessati sono stati assegnati dagli organizzatori di un corso di livello A1.1 secondo il QCER (Consiglio d'Europa & Goethe-Institut Inter Nationes ecc., 2001).

4.2 Intervento

Lo studio prevedeva interventi in diversi aspetti dell'organizzazione dei corsi. Questo approccio è stato ampiamente testato durante precedenti corsi intensivi per poi essere ottimizzato. Nella fase d'intervento, l'attuazione si è svolta mediante workshop introduttivi destinati agli organizzatori dei corsi e agli insegnanti nonché mediante visite alle classi, colloqui, linee guida e registri d'insegnamento. È stata effettuata un'analisi del materiale didattico, quest'ultimo forniva informazioni dettagliate su molti aspetti. Ciò è stato fatto sia per il gruppo d'intervento sia per quello di controllo, benché in minor misura nel secondo.

L'insegnamento per ogni coppia di classi del GI e al GC è per molti aspetti simile: il manuale di partenza (e quindi anche, per esempio, i temi) è identico, così come sono uguali il numero di ore di lezione e l'infrastruttura. Anche la durata dei compiti a casa è stata decisa di comune accordo fra gli insegnanti. Quello che differenzia i due concetti sono in parte i materiali utilizzati, ma soprattutto il punto focale dell'insegnamento.

Gli insegnanti avevano l'obbligo di seguire le istruzioni relative al concetto anche quando completavano le lezioni con attività proprie. I responsabili del progetto si sono impegnati per mettere a disposizione una quantità sufficiente di materiale didattico conforme al concetto, in modo da limitare al massimo le aggiunte di ulteriori materiali didattici.

È stato dedicato molto tempo ad adattare e ristampare i manuali per il gruppo d'intervento. Determinate parti sono state eliminate, modificate o completate, affinché il lavoro con i *chunk* diventasse preponderante. Le modifiche rispetto al concetto e ai manuali del GC possono essere riassunte come segue:

- introduzione comprendente 20-30 minuti di esercitazione alla fluenza, in cui viene rimangiato in quantità uguali quanto appreso il giorno prima e quanto appreso in precedenza. L'importante è che i mezzi linguistici utilizzati nell'esercitazione alla fluenza siano noti e che siano stati usati ripetutamente a lezione. Spesso, nell'esercitazione alla fluenza, si tratta semplicemente di ripetere le attività del manuale o del materiale di riferimento, ma con limiti di tempo.⁹ Una forma di esercizio molto versatile che viene consigliata è il

9 La fase di valutazione iniziale ha dimostrato che in determinate attività (p.es. esercizi di dialogo) mettere gli studenti sotto pressione riducendo il tempo concesso, conduce a stress e a una percezione negativa dell'esercizio. Non sono quindi state richieste attività basate principalmente sul tempo come quella suggerita da Nation, in cui uno stesso contributo orale viene ripetuto più volte riducendo il tempo a disposizione.

pointing: si parte da una parola chiave, una *flashcard* o un'immagine per creare enunciati di lunghezza e in quantità variabili. L'esercitazione alla fluenza comprende anche attività scritte, p.es. partecipare in gruppo a un gioco nel quale si cerca di trovare un massimo di frasi guardando un'immagine.

- sostituzione delle liste di vocaboli e delle tabelle di grammatica con un glossario personale (*lexical notebook*), come proposto da Lewis (1997). Il quadernetto viene riempito via via durante le lezioni e contiene numerosi enunciati personalizzati in funzione delle esigenze comunicative di ognuno.
- completamento del manuale con dialoghi riguardanti la vita quotidiana (accorgimento applicato in un solo manuale vista la quantità ridotta del materiale didattico esistente). I dialoghi riguardanti la vita quotidiana vengono introdotti anche nel materiale del GC, ma con una didattica diversa. Nel GI ci si concentra infatti sui *turn* e sui *chunk*. I dialoghi vengono messi a disposizione in forma schematizzata – anche come schede pieghevoli – e provvisti di campi lasciati in bianco.
- adattamento sistematico dello svolgimento degli esercizi. Mentre il GC lavora con vocaboli o morfemi, il GI lavora con *chunk*, sulla base di tipologie di esercizi raccomandate da Lewis (1997):
 - categorizzare, p.es. creare, a partire da frammenti, *chunk* relativi alla comunicazione
 - riordinare, p.es. conversazioni “ingarbugliate”
 - completare, p.es. dettato con spazi bianchi in cui vanno inseriti *chunk* interi o si costituiscono *chunk* rilevanti. Con il passare del tempo, quest'ultimo tipo di esercizio si sviluppa fino a dover riformulare intere battute di dialoghi della vita quotidiana
- sostituzione degli obiettivi di apprendimento fortemente improntati alla grammatica con strumenti analoghi, ma imperniati sui *chunk*

Le caratteristiche peculiari inerenti allo svolgimento degli esercizi nei manuali originali – e quindi anche nell'insegnamento del GC – sono stati in ampia misura eliminati dai materiali per il GI e sostituiti conformemente al concetto d'intervento. Più precisamente, si è intervenuto sulle attività volte a esercitare regole di grammatica, attività in cui vanno individuati e classificati fenomeni grammaticali, esercizi per l'impiego della morfologia e della sintassi, esercizi di conversione di forme grammaticali.

Sui materiali per il GC si è intervenuto solo in minima parte. Si è chiesto agli insegnanti di tenere lezioni che corrispondessero per quanto possibile alla “logica” dei manuali secondo l'interpretazione del concetto per il GC. In pratica, si trattava di integrare meno del solito materiali e idee propri. In tal modo, le lezioni del GC erano più chiaramente definite e più simili tra loro nelle varie classi a prescindere dall'insegnante.

4.3 Raccolta dei dati

Bisogna distinguere tre tipi di raccolta dei dati: 1) la raccolta dei dati riguardanti i partecipanti sulla base delle informazioni dell'organizzatore del corso; 2) i test (d'entrata e finali) orali e scritti; 3) i colloqui con gli insegnanti.

Dati riguardanti i partecipanti

Tutti gli organizzatori raccolgono una serie di dati a scopo amministrativo e per assegnare gli studenti a corsi di livello adeguato. Per la costituzione dei gruppi partecipanti al nostro studio, erano rilevanti soprattutto le informazioni sulla lingua, sulla provenienza culturale e sul livello di formazione. Informazioni più approfondite, per esempio sull'abitudine o le capacità di apprendimento ricavate da precedenti corsi, erano disponibili soltanto in alcuni casi, per cui non è stato possibile prenderle in considerazione né per la creazione dei gruppi né per l'analisi dei risultati (come variabili che spiegassero differenze nei progressi compiuti).

Test

In tutte le parti dei test sono state valutate soltanto le conoscenze e le capacità che i partecipanti potevano aver acquisito durante i corsi. Eventuali influssi della sfera privata sono stati ridotti al minimo. Sebbene fosse stata redatta una versione specifica dei test per ciascuno dei due manuali, la maggior parte degli elementi era identica. Il test finale consisteva sia in elementi che erano già apparsi nel test d'entrata sia in elementi riguardanti la seconda parte del corso. Le corrispondenze nella struttura dei diversi test erano necessarie in modo da poter rappresentare, per ciascun tipo di competenza, tutti i risultati su un'unica scala.

I test sono stati condotti sul posto da responsabili formati. Il test scritto consisteva in una prova di gruppo durata circa 80 minuti. Quello orale è stato svolto individualmente al computer, sotto la supervisione di un responsabile. Il primo test orale è durato all'incirca 30-40 minuti per partecipante, il secondo circa 35-45 minuti.

La seguente tabella presenta una panoramica dei campi (costrutti) e degli strumenti applicati e valutati nei test d'entrata e finali.

10 Il C-test è un formato testuale con spazi bianchi in cui, dalla seconda frase in poi, una parola su due è scritta solo per metà e la seconda parte delle parole è da completare (team del C-test, s.d.).

11 Nella nostra versione degli esercizi di *elicited imitation* (Erlam, 2009) al computer (con il programma PsychoPy, Peirce, 2009), i partecipanti ascoltavano una parola o una sequenza, dopodiché dovevano leggere una cifra che appariva sullo schermo. Dopo un determinato segnale acustico, dovevano ripetere la parola o sequenza. L'interferenza del numero ha la funzione di evitare che quanto udito venga copiato dalla memoria sensoriale (Shiffrin, 2003) senza essere elaborato.

12 La *speech rate* e le altre due misure sono indicatori della fluenza di enunciati orali (*utterance fluency*) (Segalowitz, 2010). Poiché la *speech rate* tiene conto della velocità di eloquio e della lunghezza delle pause, è una misura tecnica della fluenza che corrisponde bene alla fluenza percepita. Ciò non è invece il caso per la sola velocità di eloquio (*articulation rate*) (Kormos & Dénes, 2004; de Jong, Steinel, Florijn, Schoonen & Hulstijn, 2012).

13 È stata interpretata come pausa silenziosa un'assenza di articolazione per più di 250 millisecondi.

Campo / costrutto valutato	Tipo di compito [aspetto codificato]
Test scritto	
1a) Competenza scritta generale (focus sui contenuti)	C-test ¹⁰ con 4 test di 15 spazi bianchi ciascuno [inserimento di una parola perlomeno riconoscibile come quella corretta]
1b) Competenza scritta generale (focus sui contenuti e sulla forma)	C-test con 4 test di 15 spazi bianchi ciascuno [inserimento della parola completa corretta]
2) Grammatica: morfosintassi (potenzialmente sulla base di regole)	Creare frasi a partire da elementi Varianti di una stessa frase
Grammatica: morfologia (potenzialmente sulla base di regole)	Fornire il plurale
lessico-grammatica: preposizioni (potenzialmente sulla base di regole)	Completare frasi
Test orale	
3a) Comprensione e produzione orali (focus sui contenuti)	<i>Elicited imitation</i> ¹¹ [numero di elementi di contenuto riprodotti] <i>Elicited imitation</i> [numero di radici riprodotte]
3b) Comprensione e produzione orali (focus sui contenuti e sulla forma)	<i>Elicited imitation</i> [numero di parole integralmente riprodotte] <i>Elicited imitation</i> [numero di <i>chunk</i> integralmente o sostanzialmente riprodotti]
4) Qualità: formulare degli enunciati orali (<i>chunking</i>)	<i>Elicited imitation</i> [completezza e precisione dei <i>chunk</i>]
5) Grammatica: morfologia (prevalentemente intuitiva)	<i>Elicited imitation</i> [correttezza della morfosintassi]
6) Fluenza orale (fluenza cognitiva)	Produzione orale assistita mediante immagini e registrazioni audio quali esempi [a. <i>speech rate</i> ¹² – numero di sillabe al secondo, in un determinato segmento linguistico, comprese tutte le pause ¹³] [b. <i>pruned speech rate</i> – numero di sillabe al secondo, in un determinato segmento linguistico, (comprese le pause vuote e piene), escluse le sillabe delle pause piene, ripetizioni e autocorrezioni] [c. <i>articulation rate</i> – numero di sillabe al secondo, solo nel segmento linguistico con articolazione (dunque escluse le pause vuote)]

Tabella 1: Panoramica dei costrutti e strumenti dei test d'entrata e finale

Colloqui con gli insegnanti

La collaborazione con gli insegnanti si è conclusa con un colloquio tenutosi dopo la fine della fase d'intervento. I colloqui con gli insegnanti del GI e del GC divergevano soltanto rispetto alle domande relative al concetto. Sono stati trattati due ambiti principali: 1) il concetto; 2) i partecipanti. I colloqui sono durati tra 30 e 100 minuti.

I sedici colloqui sono stati trascritti e valutati mediante un'analisi dei contenuti (Mayring, 2008) per trarne i punti di vista rilevanti per lo studio.

- i. Attuazione del concetto
 - a. Confronto con i metodi d'insegnamento usuali
 - b. Attuazione concreta del concetto
 - c. Fattori d'influsso
- ii. Dinamica d'insegnamento con i singoli partecipanti
- iii. Vantaggi e inconvenienti del concetto
 - a. Opinione dell'insegnante
 - b. Opinione dei partecipanti (secondo l'insegnante)

Trattamento dei risultati dei test

Le prestazioni scritte e orali dei partecipanti sono state analizzate da una coppia di codificatori sulla base di precise direttive scritte.

Il lavoro di determinazione dell'indice di fluenza è stato particolarmente difficile a causa delle sequenze monologiche. Le registrazioni sono state segmentate, categorizzate e contate con il programma informatico *Praat* (Boersma & Weenink, n.d.), ma in parte anche manualmente. Questo lavoro ha poi permesso di calcolare gli indici.

5 Risultati

5.1 Risultati dei test

I risultati delle parti scritte dei test e nei test di *elicited imitation* (per un totale di quattro varianti di test) sono stati trattati secondo il modello di Rasch.¹⁴ Grazie a questo modello i risultati delle quattro tipologie di test (due manuali, due periodi di tempo diversi) sono stati ridotti alla stessa scala e sono dunque paragonabili.¹⁵ I dati sulla fluenza orale si sono invece potuti usare direttamente, poiché riguardavano in entrambe le occasioni lo stesso compito di produzione.

La valutazione statistica si basa su un'analisi di covarianza unifattoriale, ANCOVA, (appartenenza al GI *oppure* al GC)¹⁶ con rilevamento ripetuto (test nei due momenti t1 e t2) e attuato con modelli misti (*mixed model*).

Per la valutazione delle varie parti dei test è stato necessario tener conto di un numero leggermente diverso di partecipanti (N). 153 partecipanti avevano iniziato a seguire la seconda parte del corso; di 141 persone si sono potuti raccogliere dati per almeno una parte dei test e di 129 persone per tutte le parti dei test. Non c'è un nesso significativo tra i risultati del test d'entrata e l'abbandono del corso.¹⁷ La tabella 2 presenta una panoramica di questi dati. La mancanza di dati si spiega soprattutto per le assenze dei partecipanti dovute a ragioni organizzative per determinate parti del test.

14 Le analisi statistiche sono state condotte con pacchetti dell'ambiente di software R (Chambers, 2008; Bates, Mächler, Bolker & Walker, 2015; Kiefer, Robitzsch & Wu, 2015; Lenth, 2015; Halekoh & Højsgaard, 2014).

15 I risultati sono stati espressi in forma di WLE (Warm, 1989). Le scale WLE sopravvalutano leggermente la variabilità reale dei dati, ma per semplificare abbiamo considerato i WLE come valori osservati.

16 Diversamente da quanto accade con l'analisi della varianza, vengono incluse anche variabili (covariate) di cui è nota la capacità di influire sulle differenze individuali. Nel nostro caso, si è trattato delle origini delle persone, della durata del soggiorno in Svizzera, del livello d'istruzione e del sesso. Vi si aggiunge la variabile "coppia di classi", che consente di tenere conto del fatto che, a partire da un gruppo di persone relativamente grande, gli organizzatori dei corsi hanno creato una classe d'intervento e una classe di controllo, le quali hanno poi seguito le rispettive lezioni in condizioni molto simili.

17 Ciò è stato verificato mediante correlazioni di Pearson tra le variabili "abbandono del corso – sì o no" e gli esiti dei test nel periodo t1.

	GI	GC	Totale
Hanno cominciato la 2 ^a parte del corso	73	80	153
Hanno portato a termine la 2 ^a parte del corso (e almeno una parte di test)	66	75	141
Donne/uomini alla fine della 2 ^a parte del corso	34/32	34/41	68/73
Anni di scolarizzazione: media (minimo-massimo)	8.8 (0-17)	9.3 (0-16)	9.1 (0-17)
Età: media (minima-massima)	33.2 (16-51)	32.3 (17-57)	32.7 (16-57)
Hanno completato tutte le parti dei test e tutti i test parziali del t2; dati utilizzabili	63	66	129
Hanno completato la parte scritta del t2	66	73	139
Hanno completato la parte orale I (<i>elicited imitation</i>) del t2	63	68	131
Hanno completato la parte orale II (produzione/fluenza) del t2	64	71	135

Tabella 2: Panoramica dei partecipanti ai corsi e dei dati dei test

Nella tabella 3 è rappresentata una panoramica dei risultati dei test. I risultati rappresentano gli undici ambiti analizzati nelle tre parti dei test. Sono indicati i valori medi per ciascun gruppo e ciascun periodo, tra parentesi l'errore standard. Diversi test di significatività verificano quanto i risultati possano essere generalizzati in termini statistici. Per ciascun ambito, è valutato a) il progresso da t1 a t2 nel GI e nel GC; b) la differenza di progresso tra il GI e il GC. Sono fondamentali per il presente studio i dati sul confronto del progresso nell'ultima colonna (zigrinata scura).¹⁸

18 Per ognuna delle parti orali dei test, sono indicati due diversi valori p. Il primo di questi valori (p) include anche le prestazioni che sono state codificate e valutate più volte per diversi fenomeni, il secondo (p_{adj} , in cui H sta per "correzione di Holm"), tiene conto dell'usuale livello globale di confidenza del 95% per ciascuna parte di test. Ne consegue che i valori p con correzione di Holm sono meno sovente significativi.

Ambito analizzato	Gruppo d'intervento (GI)			Gruppo di controllo (GC)			GI vs. GC
	Test d'entrata (t1)	Test finale (t2)	Significatività del progresso ¹⁹	Test d'entrata (t1)	Test finale (t2)	Significatività del progresso	Differenza di progresso; significatività ²⁰
Parte scritta N = 139, n (GI) = 66, n (GC) = 73							
1a) Competenza scritta generale: C-test scritto / ricomposizione completa	0.064 (0.126)	0.434 (0.126)	t = 5.13 p < 0.0001	0.134 (0.113)	0.540 (0.113)	t = 5.92 p < 0.0001	F = 0.13 ²¹ p = 0.719
1b) Competenza scritta generale: C-test scritto / parola corretta	0.105 (0.081)	0.497 (0.081)	t = 4.22 p = 0.0003	0.077 (0.074)	0.448 (0.074)	t = 4.19 p = 0.0003	F = 0.03 p = 0.864
2) Correttezza grammaticale nello scritto (potenzialmente sulla base di regole: esercizi di applicazione della morfologia e della sintassi)	0.017 (0.181)	0.469 (0.181)	t = 3.79 p = 0.0013	0.386 (0.163)	0.829 (0.163)	t = 3.91 p = 0.0008	F = 0.003 p = 0.953
Parte orale I: elicited imitation N = 131, n (GI) = 63, n (GC) = 68							
3a) Comprensione e riproduzione della lingua orale: numero di elementi di contenuto riprodotti	-0.143 (0.131)	0.236 (0.131)	t = 5.74 p < 0.0001	-0.159 (0.119)	0.301 (0.119)	t = 7.24 p < 0.0001	F = 0.776 p = 0.380 p _H = n.s.
3b) Comprensione e riproduzione della lingua orale: numero di parole riprodotte integralmente	0.055 (0.110)	0.379 (0.110)	t = 5.89 p < 0.0001	0.025 (0.099)	0.495 (0.099)	t = 8.88 p < 0.0001	F = 3.64 p = 0.059 p _H = n.s.
3c) Comprensione e riproduzione della lingua orale: numero di radici riprodotte	-0.126 (0.135)	0.229 (0.135)	t = 5.42 p < 0.0001	-0.164 (0.122)	0.276 (0.122)	t = 6.99 p < 0.0001	F = 0.89 p = 0.348 p _H = n.s.
4) Caratteristica formulare degli enunciati orali: <i>chunking</i>	-0.099 (0.119)	0.098 (0.119)	t = 3.27 p = 0.007	-0.160 (0.108)	0.205 (0.108)	t = 6.31 p < 0.0001	F = 4.075 p = 0.046 p _H = n.s.
5) Correttezza grammaticale (piuttosto) intuitiva: correttezza della morfosintassi	-0.029 (0.142)	0.248 (0.142)	t = 3.50 p = 0.0035	-0.196 (0.129)	0.455 (0.129)	t = 8.545 p < 0.0001	F = 11.566 p = 0.0009 p _H = 0.0045
Parte orale II: produzione orale N = 135, n (GI) = 64, n (GC) = 71							
6a) Fluenza orale: numero di sillabe per tempo di registrazione in secondi; <i>speech rate</i>	1.490 (0.054)	1.610 (0.054)	t = 2.73 p = 0.035	1.527 (0.049)	1.516 (0.049)	t = -0.29 p = 0.99	F = 4.765 p = 0.031 p _H = 0.093
6b) Fluenza orale: numero di sillabe per tempo di registrazione in secondi, senza correzione di sillabe, ripetizioni e pause piene; <i>pruned speech rate</i>	1.190 (0.053)	1.331 (0.053)	t = 3.27 p = 0.0075	1.214 (0.048)	1.230 (0.048)	t = 0.39 p = 0.98	F = 4.42 p = 0.037 p _H = n.s.
6c) Fluenza orale: numero di sillabe per tempo di registrazione escluse pause vuote; <i>articulation rate</i>	3.006 (0.057)	3.075 (0.057)	t = 1.35 p = 0.53	3.118 (0.051)	3.086 (0.051)	t = -0.65 p = 0.92	F = 2.02 p = 0.157 p _H = n.s.

Tabella 3: Panoramica dei risultati dei test

Se si prende in considerazione la correzione di Holm (valori p_H), una sola variabile, nell'ambito 5 (correttezza della morfosintassi nella parte orale, *elicited imitation*), segna una differenza tra il progresso del GI e quello del GC. In quest'ambito il progresso nel GC è mediamente superiore rispetto a quello del GI. Escludendo il fattore dei test molteplici (valore p), sono significativi o al limite del significativo i seguenti confronti di progresso (in corsivo):

- a favore del GC: 3b) numero di parole riprodotte integralmente e 4) *chunking*
- a favore del GI: 6a) *speech rate* e 6b) *pruned speech rate*

In tre ambiti valutati su cinque per l'*elicited imitation*, appare quindi un vantaggio per il gruppo di controllo, mentre negli altri due i progressi dei due gruppi sono paragonabili. È interessante un confronto dal punto di vista dei contenuti: il GC presenta un progresso maggiore quando il *focus* è sulla correttezza e la completezza (ambiti 3b, 4 e soprattutto 5). Quando invece gli aspetti formali assumono un'importanza minore (3a, 3c), le differenze tra i due gruppi sono minime.

Il progresso dal test d'entrata a quello finale è quasi sempre significativo *in seno* a ciascuno dei due gruppi. Fanno eccezione (vedi le quattro celle zigrinate chiare) le valutazioni della fluenza. Nel GC, tra t_1 e t_2 non c'è stato alcun cambiamento per le tre misure di fluenza, mentre nel GI è rimasta invariata soltanto l'*articulation rate*, la misura della velocità d'eloquio senza tener conto delle pause. I valori medi della *speech rate* e della *pruned speech rate* sono invece aumentati in modo significativo, rispettivamente dell'8% e del 12%.

19 Per ognuna delle parti orali dei test, sono indicati due diversi valori p . Il primo di questi valori (p) include anche le prestazioni che sono state codificate e valutate più volte per diversi fenomeni, il secondo (p_H , in cui H sta per "correzione di Holm"), tiene conto dell'usuale livello globale di confidenza del 95% per ciascuna parte di test. Ne consegue che i valori p con correzione di Holm sono meno sovente significativi.

20 Valori p del t-test per quanto riguarda il progresso per gruppo con correzione secondo Tukey per test multipli.

21 Valore p dell'F-test con correzione per *mixed-models* (approssimazione Kenward-Roger).

5.2 Risultati dei colloqui con gli insegnanti

Di seguito vengono presentate, nello stesso ordine come al punto 4.3, le importanti conclusioni ottenute in seguito ai colloqui con gli insegnanti alla fine del progetto.

Attuazione del concetto da parte degli insegnanti

Come summenzionato, allo scopo di garantire la comparabilità dell'insegnamento in tutte le classi dei due gruppi, è stato elaborato, introdotto e redatto sotto forma di linee guida un concetto specifico per ciascun gruppo. L'attuazione è inoltre stata accompagnata da consulenze, da visite in classe e dalla consultazione dei registri degli insegnanti. Dai colloqui emerge che le lezioni sono state in gran parte improntate al concetto come auspicato. I materiali personali e le attività non in linea con il concetto sembrano avere avuto un'importanza del tutto marginale. Nelle classi del GI, l'elemento nuovo dell'esercitazione alla fluenza di 20-30 minuti all'inizio di ogni lezione è stato in ampia misura attuato come richiesto, anche se, soprattutto in un primo momento, non è stato semplice.

Accettazione dei concetti da parte dei singoli partecipanti e delle classi

Il concetto d'intervento è stato accolto in vario modo dalle classi. Le modalità più dinamiche e sociali con le quali venivano svolti alcuni esercizi orali (di ripetizione) hanno suscitato resistenza presso alcuni partecipanti (p.es. quelli più anziani) e entusiasmo presso altri. Un'insegnante reputa che la maggiore interazione abbia aiutato a creare un clima di classe particolarmente costruttivo.

Sembra che il concetto del GI sia stato meglio accettato dai partecipanti più deboli e meno abituati all'apprendimento scolastico che da quelli con maggiore familiarità con l'ambiente scolastico e da quelli che desideravano migliorare la correttezza della lingua. Per i soggetti della prima categoria, l'assenza dell'insegnamento della grammatica è stato sentito come un sollievo. Gli studenti invece più abituati a un concetto scolastico hanno talvolta espresso dubbi sull'utilità di esercitare la fluenza se gli errori non vengono o vengono solo in minima parte corretti. Queste persone hanno inoltre espresso il desiderio di ricevere nozioni formali di grammatica o materiale per studiarla. Hanno tuttavia riconosciuto l'utilità didattica degli esercizi alla fluenza come opportunità di ripetere regolarmente quanto appreso.

Nelle classi del GC, vi sono state reazioni negative soprattutto in merito all'apprendimento a memoria di liste di vocaboli come previsto dal manuale.

Ambedue i gruppi hanno reagito molto positivamente ai dialoghi della vita quotidiana su CD che sono stati distribuiti come materiale supplementare. Ne è stata approntata una versione con una didattica specifica a ciascun concetto.

Valutazione dei vantaggi e degli inconvenienti dei concetti da parte degli insegnanti

Diversi insegnanti di ambedue i gruppi hanno lamentato i limiti dovuti all'obbligo di adottare i concetti e i materiali prestabiliti. In particolar modo quelli che hanno lavorato con il manuale meno specificamente orientato al gruppo mirato avrebbero desiderato poter tralasciare alcuni elementi didattici di tipo tematico, lessicale e grammaticale, e eventualmente sostituirli con i propri. Gli insegnanti del GI hanno invece osservato che l'attuazione del concetto d'intervento non sarebbe stata possibile senza gli appositi materiali didattici che sono stati loro forniti.

Una parte degli insegnanti del GI si è espressa favorevolmente sulle varie componenti del concetto d'intervento, un'altra parte in modo più critico. È stata apprezzata l'importanza data, soprattutto grazie all'esercitazione alla fluenza, all'aspetto orale. Gli insegnanti hanno espresso perplessità sulla quantità ridotta di materiale disponibile per gli esercizi scritti. L'individualizzazione nell'esercitazione alla fluenza è stata percepita come un problema, perché i partecipanti tendono a fare errori nel parlato e a ripeterli frequentemente senza che l'insegnante possa intervenire utilmente.

Gli esercizi con i *chunk* che per il GI hanno sostituito quelli incentrati sulla grammatica hanno ricevuto un giudizio tutto sommato positivo. Al contempo, è stata problematizzata l'assenza di un insegnamento esplicito della grammatica. Da un lato, chi aveva familiarità con l'apprendimento non ha ricevuto nozioni né materiale di grammatica, dall'altro agli insegnanti è mancata la possibilità di usare la metalingua grammaticale per correggere gli errori e fornire spiegazioni.

Il glossario personale (*lexical notebook*) è stato generalmente giudicato come uno strumento utile. È tuttavia stato osservato che le correzioni necessarie richiedono molto tempo. Infatti, i discenti inseriscono da soli i vocaboli nel quaderno e l'insegnante deve correggerli, poiché è sulla base del glossario che sono previste le ripetizioni individuali. Inoltre, per chi non ha molta familiarità con l'apprendimento scolastico occorrono troppe istruzioni per riuscire a lavorare con questo strumento. Gli elementi individualizzati come il glossario e le schede con i dialoghi della vita quotidiana sono considerati favorevolmente perché stimolano l'autonomia nell'apprendimento.

Al contrario del concetto d'intervento, quello del gruppo di controllo (ossia l'insegnamento secondo le proposte contenute nel manuale per i docenti) è stato considerato adatto per chi ha familiarità o facilità nell'apprendimento. Tale valutazione non è però condivisa allo stesso modo da tutti. Qualcuno ha rilevato che il concetto contiene troppa grammatica, troppi concetti grammaticali e un lavoro lessicale eccessivamente

esplicito. Si è però anche rilevata l'importanza di questo lavoro intensivo sul lessico. Secondo un'insegnante, il concetto del GC dovrebbe essere mantenuto per sensibilizzare all'apprendimento scolastico anche chi non vi è abituato.

6 Discussione

Dai risultati dello studio emerge complessivamente un quadro ricco, sia sull'utilizzo dei due concetti che sull'evoluzione delle competenze linguistiche dei partecipanti.

Dai colloqui, si sono potuti trarre molti spunti per rispondere alle prime tre domande della nostra ricerca (cfr. cap. 3). Il nuovo concetto d'insegnamento e i relativi materiali (domanda 1) sono stati sviluppati, attuati e introdotti con buon esito. Durante i colloqui, qualche insegnante di ambo i gruppi ha lamentato il carattere fortemente pilotato dell'insegnamento nella fase d'intervento. Gli insegnanti sarebbero voluti intervenire con materiali propri, orientati alle esigenze dei partecipanti. Il pilotaggio è stato tuttavia necessario in entrambi i gruppi per gli scopi della ricerca, poiché altrimenti sarebbe stato arduo ricondurre i risultati all'influsso dei due concetti.

L'accettazione del concetto d'intervento da parte degli insegnanti e dei partecipanti (domanda 2) è variabile e va valutata in dettaglio. Il concetto esclude alcuni elementi ai quali si è abituati nel contesto dei corsi di lingue straniere (in particolar modo il lavoro grammaticale esplicito) e pone un forte accento su altri (soprattutto l'esercitazione quotidiana alla fluenza e attività con i *chunk* per creare nessi funzionalmente logici, ma non esplicitamente analizzati). Per questo motivo non corrisponde alle aspettative dei discenti abituati ai corsi di lingua tradizionali e di quelli che frequentano il corso con l'obiettivo di migliorare la correttezza della lingua. Il concetto d'intervento, essendo diverso dalla prassi abituale, costituisce inoltre una sfida per gli insegnanti. Un problema sentito al contempo dagli insegnanti e dai partecipanti è che l'esercitazione alla fluenza induce l'automatizzazione di una lingua che non si padroneggia ancora abbastanza. In realtà, l'esercitazione alla fluenza ha come scopo l'utilizzo della lingua già nota, ossia quella che si padroneggia relativamente bene (Nation & Meara, 2002). A quanto pare, i partecipanti disponevano di una quantità di materiale linguistico insufficiente rispetto alle prestazioni richieste e di conseguenza facevano numerosi errori. Un altro problema sentito è stata l'assenza di un insegnamento esplicito della grammatica. I partecipanti con una certa abitudine all'apprendimento hanno chiesto di ricevere nozioni grammaticali e/o materiale sull'argomento. Gli insegnanti hanno fatto fatica a rinunciare alla terminologia grammaticale quando si trattava di correggere o dare riscontri. I partecipanti più deboli o meno abituati all'apprendimento scolastico sembrano invece aver apprezzato il fatto di non dover addossarsi il peso di uno studio esplicito della grammatica. Sulla base delle valutazioni degli insegnanti, possiamo rispondere in maniera perlopiù positiva alla terza domanda che ci siamo posti: il concetto d'intervento è adatto a chi ha poca familiarità con l'insegnamento scolastico. Una delle difficoltà didattiche più frequenti nei corsi di base è l'eterogeneità del pubblico. Lo si è visto nel nostro studio in particolare in relazione con le diverse reazioni riguardo all'insegna-

mento esplicito della grammatica. Si potrebbe quindi concludere che il concetto d'intervento è più adatto ai discenti poco abituati all'apprendimento scolastico oppure che conviene perlomeno affiancare tale concetto a elementi didattici convenzionali, in particolare con nozioni di grammatica (che offre una terminologia utile per le correzioni e agevola lo studio individuale). L'esercitazione specifica alla fluenza proposta da Nation & Meara (2002) potrebbe, al contrario, confluire nel concetto convenzionale: da una parte perché con un maggiore *focus* sulla correttezza il rischio di esercitare errori si ridurrebbe, dall'altra perché, sebbene la fluenza sia considerata un fattore importante nel processo di apprendimento delle lingue, un suo sviluppo mirato è solitamente ignorato nei modelli classici d'insegnamento.

I risultati dei test corrispondono parzialmente alle attese (ottimiste) che abbiamo formulato seguendo quanto viene affermato nella letteratura scientifica sui benefici di un approccio basato sui *chunk* (vedi punto 3).

Quanto atteso si è verificato in tre ambiti:

- i progressi nella fluenza sono maggiori nel GI
- come previsto, i progressi lessicali (sul piano dei significati) e nella capacità di comunicare contenuti non differiscono da un gruppo all'altro
- i progressi nella correttezza dello scritto sono analoghi nei due gruppi

In due ambiti, le ipotesi non si sono verificate:

- i progressi nell'impiego di *chunk* più completi e corretti sono risultati minori nel GI che nel GC
- i progressi nella correttezza grammaticale nell'uso orale sono stati maggiori nel GC

È interessante interpretare questi risultati nel loro insieme, confrontandoli con quelli dei colloqui con gli insegnanti.

Rimane aperto un interrogativo: in che misura il quadro generale dei risultati dei test corrisponde all'osservazione dei docenti sul fatto che nell'esercitazione alla fluenza (ossia attività frequenti focalizzate sulla ripetizione) si esercitassero sovente errori linguistici. Ripetere elementi scorretti conduce a memorizzare e a ricorrere a *chunk* o *construction* non conformi alla lingua bersaglio? I risultati dei test non sembrano contraddire questa ipotesi: gli enunciati orali del GI sono mediamente meno corretti di quelli del

GC. Sembra indicarlo il confronto nei tre ambiti in cui interviene una valutazione della correttezza con l'*elicited imitation*, ossia: correttezza grammaticale, precisione lessicale e *chunking* corretto. Al contempo, la fluency orale è migliorata solo nel GI e questo non in termini di maggiore velocità di articolazione (*articulation rate*), ma, come evidenziano i risultati della *speech rate* e della *pruned speech rate*, di una quantità minore di pause (silenzi, riempitivi, ripetizioni e correzioni) durante l'elocuzione. Pause più lunghe (o un'elocuzione meno produttiva) denotano una maggiore esigenza di pianificazione da parte dei partecipanti del GC. I partecipanti del GI sembrano invece essere in grado di ricorrere maggiormente a modelli pronti, un indicatore dell'avvenuta proceduralizzazione della lingua imparata (de Jong & Perfetti, 2011) o dell'esistenza di *construction* che è possibile impiegare prontamente. I risultati sulla correttezza nello scritto, pressoché identici nei due gruppi, non contraddicono questa interpretazione. È possibile che dato il poco tempo a disposizione per l'*elicited imitation*, si sia manifestata la natura irrefrenabile (*ballistic*) della lingua automatizzata (Segalowitz & Hulstijn, 2009), mentre nella prova scritta è emerso l'effetto correttivo principalmente perché il locutore aveva più tempo a disposizione per ricorrere ad altro sapere. Rimane da chiedersi di quali forme di sapere si sia trattato: sono stati confrontati vari *chunk* acquisiti oppure si è ricorso a conoscenze (regole) astratte dedotte personalmente dalla propria esperienza di apprendimento linguistico?

Gli esiti delle parti orali dei test vanno considerati anche dal punto di vista dei risultati di ricerca sulla *performance* nelle attività orali, in particolare tenendo conto dell'ipotesi del *trade-off* (Skehan, 2009). Il presupposto dell'ipotesi è che le nostre capacità cognitive siano limitate. La memoria a breve termine ci consente di mantenere l'attenzione solo fino a un certo punto. Per questo, quando si parla, sussiste una concorrenza tra le esigenze di complessità, correttezza e fluency linguistiche.²²

Si suppone che nel caso dei nostri test, viste le istruzioni chiaramente definite, la complessità non sia stata un fattore decisivo e in ogni caso sia stata uguale per ambedue i gruppi. I partecipanti hanno quindi potuto concentrare le loro risorse cognitive essenzialmente sulla correttezza e sulla fluency. Il GI sembra aver prestato più attenzione alla fluency, il GC alla correttezza. Ciò non corrisponde a quanto presupposto dai sostenitori dell'approccio lessicale (Nation & Newton, 2009), poiché essi ipotizzano che la possibilità di ricorrere a elementi interi, ossia ai *chunk*, possa migliorare non solo la

22 Il rapporto tra complessità, correttezza e fluency (*complexity, accuracy, fluency*, abbreviato in CAF) è oggetto di dibattiti tuttora in corso nella linguistica applicata e nella ricerca sui test di lingua.

fluency ma anche la correttezza. Ci si potrebbe aspettare che il fatto di avere un facile e rapido ricorso ai *chunk* liberi risorse impiegabili in operazioni di modifica e correzione qualora i *chunk* richiedessero adattamenti grammaticali. I dati tuttavia non lo confermano. È ipotizzabile che, nel nostro caso, i risultati non riflettano tanto le risorse cognitive limitate impiegate durante la produzione orale, quanto il carattere diverso delle conoscenze e capacità linguistiche sviluppatesi durante le lezioni in seno ai due gruppi. Il GI sembra disporre di un numero maggiore di *chunk* che, benché scorretti, “emergono” con forza irrefrenabile (*ballistic*) mentre il GC dispone delle conoscenze linguistiche e dell'attenzione necessaria alla produzione orale di un tedesco più corretto, ma a discapito della fluency.

Anche i nostri risultati inficiano l'ipotesi *non-interface-position* (Krashen, 1982): l'assimilazione della grammatica per un uso linguistico spontaneo sembra riuscire meglio proprio *grazie* all'apprendimento esplicito. È tuttavia necessario precisare che, nell'insegnamento al GI, in cui i fenomeni grammaticali compaiono soltanto in *chunk* usati frequentemente, si denota un progresso analogo nell'uso scritto e formulare della lingua. Sembrerebbe, dunque, che si sia verificato un apprendimento incidentale di risorse grammaticali riutilizzabili negli esercizi di grammatica scritti.

Ma come valutare il presupposto secondo il quale, a un determinato momento, le classi che hanno ricevuto un insegnamento basato sulla memorizzazione di *chunk* ricostruiscono a partire da essi l'analisi grammaticale della lingua già impiegata? Potrebbe trattarsi di un'ipotesi valida, ma dipende dal valore che diamo, da un punto di vista cognitivo, alla differenza osservata tra la correttezza orale e scritta nel GI. I progressi relativamente maggiori nello scritto vanno considerati come un segno dell'avvenuta memorizzazione di forme più corrette, che vengono però richiamate in modo meno automatico (ossia con un minor livello di *ballistic*)? In tal caso, è senz'altro possibile che avvenga un'analisi efficace delle forme scritte, soprattutto se si continua a frequentare corsi e specialmente lezioni in cui l'obiettivo è la messa a disposizione di una lingua formalmente corretta.

Non è da trascurare il fatto che uno sviluppo della fluency sia stato osservato soltanto nei discendenti del GI (vedi i risultati delle *speech rate*): esso conferma quanto anticipato da Lewis (1993) e da altri esponenti dell'approccio lessicale. La fluency è uno degli ambiti principali in cui devono svilupparsi le competenze linguistiche. Nel GC questo non sembra succedere, il che potrebbe essere un semplice sintomo di esigenze concorrenti (correttezza e fluency) al quale i limiti della memoria a breve termine non sanno far fronte. In particolare i risultati della *pruned speech rate*, che rispecchiano anche gli aspetti della pianificazione e dell'autocorrezione, lasciano supporre che per

quanto riguarda l'espressione orale il GC abbia incontrato più difficoltà del GI. In futuro, un obiettivo per gruppi di questo genere dovrebbe essere lo sviluppo, attraverso attività di ripetizione (di *chunk* o *construction* richiamabili), di una lingua più automatizzata. In ogni caso, è certo che i contatti linguistici vengono facilitati da un uso più sciolto della lingua, il che migliora le probabilità sul lungo termine che l'apprendimento prosegua anche al di fuori delle lezioni.

7 Conclusioni

Sulla scorta dei risultati ottenuti, la domanda pratica all'origine del presente studio esige una risposta articolata. Un insegnamento imperniato sui *chunk* e che rinuncia a un'analisi esplicita della grammatica (*focus on forms* e *focus on form*,²³ Long, 1991), può essere attuato con buoni esiti nei corsi di base, ma viene accolto con riserve da discenti con una certa esperienza nell'apprendimento di lingue straniere o che mirano a migliorare la correttezza. Secondo gli insegnanti, chi non ha familiarità con l'apprendimento scolastico apprezza invece questo concetto, considerandolo meno impegnativo. Gli insegnanti fanno notare che le lezioni basate sul principio d'intervento necessitano di materiali speciali e sottolineano un aspetto problematico del concetto, ossia la frequente e apparentemente inevitabile ripetizione di una variante del tedesco che non corrisponde alla lingua bersaglio. Proprio questa caratteristica può aver causato dei minori progressi nella correttezza orale del GI rispetto al GC, il quale, mediante numerose attività più convenzionali, si è potuto concentrare sulle forme linguistiche. Ma in un ambito altrettanto importante – quello della fluenza linguistica – il GI ha fatto progressi che non sono stati misurati nel GC nella fase d'intervento. Questa constatazione parla a favore delle *balanced learning opportunities* (Ellis, 2005b). Nel concetto del GI è in ogni caso necessario prestare più attenzione alla correttezza dei *chunk* automatizzati. Ciò non significa necessariamente studiare la grammatica in modo sistematico, ma concentrarsi maggiormente sulle forme linguistiche in generale, per esempio insistendo di più sulle forme corrette della lingua bersaglio che i discenti impiegano durante l'esercitazione alla fluenza. Nel concetto del GC manca invece un incoraggiamento mirato all'automatizzazione mediante frequenti attività di ripetizione.

Per quanto riguarda ulteriori ricerche scientifiche, sarebbe auspicabile un'analisi approfondita delle stesse produzioni orali che sono state usate per determinare gli indicatori di fluenza per il presente studio. L'analisi consentirebbe di approfondire le differenze di "formulazione del linguaggio" e dei processi cognitivi dei due gruppi durante la produzione. Esistono, per esempio, differenze specifiche tra i due gruppi per quanto riguarda il tipo di errori formali? Esistono differenze su come vengono prodotte le pause piene, le correzioni e le ripetizioni? Sarebbe inoltre opportuno studiare l'efflusso di un'esercitazione alla fluenza nel quadro di un insegnamento scolastico classico di lingua straniera, che solitamente non è intensivo e non comporta molti contatti linguistici extra scolastici. Bisognerebbe inoltre tener conto dei vari livelli di compe-

23 *Focus on forms*: insegnamento sistematico ed esplicito della grammatica; *focus on form*: insegnamento esplicito della grammatica basato sulle esigenze di chiarimento che emergono durante un insegnamento orientato alla comunicazione.

tenza degli studenti al fine di soppesare l'incognita legata all'estensione del repertorio di forme corrette già acquisite che potrebbe condurre a esiti diversi per quanto riguarda la correttezza formale.

Fluency and Accuracy – Fluency or Accuracy?

Results of an intervention study in low-threshold intensive courses for adult beginners of German as a second language

—
Peter Lenz, Malgorzata Barras, Fabienne Manz¹

¹ Our thanks go to Thomas Studer, the head of this research area, for his helpful suggestions on conducting the project and his input on composing this report.

Abstract

The present study explores the potential value of implementing the Lexical Approach (Lewis, 1993; 1997) and regular fluency trainings (Nation & Meara, 2002) in courses teaching German as a second language to adult beginners who, to varying degrees, have not had much previous schooling. When implementing the Lexical Approach, learning chunks (multiple-word items, partially with open slots) is crucial. Based on a quasi-experimental design with an intervention group (IG) and a control group (CG), twelve classes divided into six pairs worked their way through an entire textbook. The teaching concept and the materials used in one class of each pair implemented the intervention concept, while the concept and the materials used in the other class were mainly in line with the course design foreseen by the original materials, and thus served as a control. Major features of the CG concept include explicit grammar instruction and vocabulary work with single words. Various written and oral language skills were measured in an initial and a final test. Interviews with teachers were held to shed light on the suitability of the teaching concepts for the target group and to better understand the specific features of the teaching and learning processes in both groups.

The interviews lead to the conclusion that, in general, both concepts can be applied to good effect. In part, IG learners with experience in learning foreign languages as well as students interested in improving accuracy missed grammar instruction and exercises. Overall they viewed fluency training as a good opportunity to learn. According to the teachers, the IG concept is better suited for students who have had less formal schooling, while the CG approach meets the needs of students with a stronger educational background. Teachers in the IG sometimes missed having a grammatical metalanguage to correct errors; they furthermore found it problematic that they had few opportunities to intervene during fluency training, in which learners frequently repeated, and thus reinforced, incorrect language.

The results in the various partial skills indicate an improvement between the initial test and the final test, with a few exceptions regarding the measurements on fluency. While the two main fluency indicators reveal no improvement in the CG, the results in the IG tell another story. The improvements in *speech rate* (number of syllables per second of the total speaking time including all pauses) and the *pruned speech rate* (identical to the speech rate but excluding syllables in filled pauses, repetitions, and self-corrections) strongly indicate that learners in the IG were able to produce oral speech with less planning and self-monitoring than learners in the CG. The detailed results from another oral test section suggest that a gain in fluency is accompanied by a loss in accuracy. In terms of linguistic function and content, the oral skills in both groups reveal no differences. There are also no differences in the written section, including the area

of morphosyntax, where an advantage for CG learners was expected. The findings are discussed in terms of theoretical approaches and concepts concerning language acquisition and language production, and suggestions for the compared teaching concepts and future research are made.

1 Introduction

This study is based on the project “The influence of foreign language teaching methodology on the acquisition of language skills by adult migrants in low-threshold language courses,” which was conducted in the scope of the 2011–2015 research program at the Research Centre on Multilingualism, with additional funding from the State Secretariat for Migration (SEM). The project aimed to shed light on a question framed by experts developing the Swiss framework curriculum to promote language skills in migrants (Lenz, Andrey & Lindt-Bangerter, 2009) by means of an empirical study. The issue at hand is whether instruction in German as a second language for learners who frequently have not had much formal schooling should work with standard teaching approaches rooted in explicit, rule-based grammar instruction (*focus on forms* according to Long, 1991) and single-word learning, or whether it is preferable for such courses to emphasize the learning and automatization of unanalyzed linguistic sequences (so-called chunks). This report aims to answer this question by presenting empirical evidence from the classroom.

For the purpose of this study, a quasi-experimental design with an intervention group and a control group was chosen. Six pairs of classes were formed, in which one class was taught using a chunk-based approach while the other class received lessons focusing on teaching grammar and single-word learning, the concept that is actually foreseen by the teaching materials.

Originally, four language schools offering low-threshold courses of German as a second language (*Deutsch als Zweitsprache – DaZ*) participated in our study. Three schools with a total of twelve classes were able to complete the study as planned, while one institution left after the first phase of the project because the heterogeneity in the classes made a redistribution of the learners necessary. The project ran very smoothly thanks to the cooperation and the great effort made by all parties involved.²

The teaching materials were a key element in the implementation of the study. The *DaZ* textbooks normally used at the various schools were also used in the study. The three schools remaining in the study worked with two different textbooks.³ To accommo-

2 We would like to offer our sincere thanks to the following schools: aoz (Zurich), FORMAZIONE (Bern), Lernpunkt (Bern) and ECAP (Central Switzerland/Lucerne).

3 The following textbooks, which have been adapted for teaching German as a second language in Switzerland, were used: *Schritte plus 1. Niveau A1/1. Ausgabe Schweiz* (Niebisch, Penning-Hiemstra, Specht & Bovermann, 2011), published by Hueber Verlag and *Studio d A1. Schweizer Ausgabe* (Funk, Kuhn, Demme & Bayerlein, 2005), published by Cornelsen Verlag.

date the chunk-based approach, various elements in these textbooks required targeted adjustments or additions. The material was then printed as a brochure.⁴

Few comparable studies have been conducted thus far. To begin, little research has been done on language learning of adult migrants although language support in a migration context has gained great social significance. Another reason is that quasi-experimental interventions lasting several weeks and focusing on overall course concepts are rare. Such interventions not only create practical problems; they also pose a particular challenge from a research perspective because factors that potentially dilute the influence of intervention concepts are more difficult to contain in everyday classroom settings than in isolated interventions. On the positive side, however, findings from interventions that form an integral part of the classroom’s eco-system better allow conclusions to be drawn with regard to classroom practice.

4 Our thanks go to all participants who gave us the rights to use their materials.

2 Context

The core of the present study is a comparison between the development of language skills of beginning learners in intensive German as a second language (*DaZ*) courses that are taught using different conceptual approaches. The general goal of these courses is to improve the language ability of the course participants, thus enabling them to better participate in the social life of their host country. This study is, however, less interested in the holistic testing of language ability, on the basis of communicative tasks; rather, it focuses on various linguistic resources and thus on partial aspects of linguistic competence such as spoken fluency and grammatical accuracy. Placing priority on linguistic resources follows in the tradition of studies in applied linguistics and generates differentiated data that is easier to interpret. Moreover, this approach is justified in that the development of a linguistic repertoire, which serves as the foundation for further language use, is the foremost aim of beginning language classes.

“Fluency” in a language is the highest goal for many language learners. As a rule, this means that someone can speak a language quickly, correctly, and without difficulties (Segalowitz, 2005). At the cognitive and physiological-articulatory level, such fluency presupposes automated processes. Acquiring this complex behavior calls for attentiveness, concentration, and effort; in general, performance is initially slow and error-prone but then improves through repetition, ultimately leading to greater ease in production (Aguado, 2002). Automatization of language use, however, is not solely related to increased speed of language processes; it also implies a restructuring of memory processes (Anderson et al., 2004; Gatbonton & Segalowitz, 2005) that reduces an individual’s reliance on working memory when using language. Rapid, effortless, unconscious, and unstoppable (ballistic) use of language is what characterizes automaticity, although all characteristics do not necessarily need to be present at once (Segalowitz & Hulstijn, 2009). Indeed, Segalowitz & Hulstijn consider ballisticity to be the key criterion for automatized language. The term ballisticity is used to describe the phenomenon that the use of automatized language is relatively immune to disturbances. In connection with L2 acquisition, the ballistic nature of automatisms may be partially responsible for the persistence of non-standard linguistic forms, i.e. fossilization⁵ (Bybee, 2008).

5 Long (2005) expresses doubts as to whether permanent “fossilization” is a real cognitive phenomenon. In a critical analysis of various studies, he demonstrates that this much-feared, deficient final condition in (foreign) language acquisition is probably not often present where it has been diagnosed. Instead, non-standard linguistic characteristics primarily indicate an incomplete process of language acquisition, which can have diverse causes.

If automatization, or the automaticity of language, is a key component of language development, the question arises as to how such processes can be attained in formal foreign language learning. From the standpoint of *Usage-Based Linguistics*, natural processes of language acquisition should be activated and applied, and ultimately lead to automatic language use. The teaching and learning challenge lies in creating and further developing so-called *constructions*. The basic assumption is that language learners (L1 and other languages) derive elements and conventions of a language from their communicative experience. Once the abstraction process (proceduralization) is completed, knowledge and skills grounded in experience are then stored in the memory as constructions, which mentally assign certain linguistic forms to certain (communicative) functions (Ellis, 2005a). Constructions, then, consist of fixed (cognitive) chunks (“neuro-motor routines”) and a variable number of interchangeable elements (Bybee, 2008). A construction may be an individual word, an expression, or an entire sentence. For instance, a construction with an open slot for variations is the half-fixed expression “I would like...”. Before a construction can be applied in language use, it first must be identified as such and then automatized through repetition. As soon as enough automatized constructions are in a learner’s repertoire – especially more comprehensive examples that themselves contain smaller units – fluent language use can be achieved. In fluent language, the constructions alleviate the effort required of the working memory with its limited capacity, thus freeing up cognitive resources for other tasks such as action control (Skehan, 2009).

Bybee (2008) explores the possibilities of a curriculum that promotes the development of constructions. A central factor consists of designing the learning conditions to compensate for the low exposure to the L2 (in comparison to L1). First, corpus analysis should be applied to select relevant language sequences. Based on these sequences, key constructions can then be taught using prototypical examples. To make such examples recognizable as a pattern and to initiate acquisition, learners should be repeatedly confronted with the same as well as modified patterns in a principled manner.

Segalowitz & Hulstijn (2009) raise the point that, while linguists would recommend such or similar practices to promote automatization, a great deal of resistance must be expected in practice. Although such an approach offers plenty of repetition to foster automaticity, in the classroom it runs the risk of being viewed as decontextualized and uninteresting. By contrast, typical communicative lessons are generally perceived as relevant, but they do not foresee enough repetition to generate automaticity. The authors see a solution in concepts that are oriented on actual communicative needs and that teach versatile target-language items as part of meaningful communicative class-

room activities. These items are then repeated sufficiently often within related communication tasks.

Our concepts for the intervention and control groups define didactically divergent approaches, especially in their treatment of grammar. The control group learns and practices grammar in explicit, systematic instruction (*focus on forms*); in the intervention group, by contrast, grammar is not addressed as such. With this approach, a long-standing, heavily contended question among researchers on language acquisition comes to the fore: can learners use language knowledge that was taught (and practiced) using explicit, rule-based methods as spontaneously as they can access knowledge that they themselves have (implicitly) inferred from input in the foreign language? Particular prominence was achieved by Krashen's (1982) radical position, which answers the question with a "no" (*non-interface position*), claiming that rule-based language knowledge functions solely as a so-called monitor to control language use. Meanwhile, however, scholars have accepted the theory that complex learning processes such as language acquisition can indeed begin with explicit instruction (declarative knowledge) and then lead to the formation of internalized procedural knowledge through practice (cf. Anderson's well-known model of cognitive development, Anderson et al., 2004). Moreover, a large meta-study on the success of various teaching methods (Norris & Ortega, 2001) provides strong indications that language teaching with a systematic or specific focus on linguistic phenomena is more successful than purely implicit teaching approaches. Indeed, it appears that the best results are generated by teaching methods that also build on explicit knowledge of the language being learned. A recent summary report by Spada & Tomita (2010), which includes more open test forms and more results on implicit instruction than the earlier study, confirms the general advantage of instruction with explicit sections over a purely implicit teaching approach. Although our intervention concept includes numerous types of exercise with a focus on specific linguistic phenomena, such phenomena are not explicitly addressed. As such, we have maintained a clear differentiation to the concept applied in the control group, which foresees a relatively large amount of explicit grammar instruction.

To implement the intervention concept, we primarily relied on an approach from the area of foreign language teaching: Lewis's *Lexical Approach*,⁶ which has garnered a great deal of attention since the 1990s (cf. Lewis 1993; 1997; 2000). We complemented this approach with suggestions by Nation (2001) on training fluency.

Lewis bases his approach mainly on corpus linguistic findings concerning the frequency

6 Willis (1990) und Nattinger & DeCarrico (1992) have conducted other, lesser known attempts to implement findings on chunks in teaching practices.

of more or less fixed connections between lexical elements. According to Lewis, lexical elements provide the key to grammar: "Language consists of grammaticalized lexis, not lexicalized grammar" (Lewis, 1993: 34). He therefore recommends taking existing multiple-word items as the starting point in language learning because, as he contends, language consists largely of such prefabricated items that cannot be arbitrarily manipulated based on grammatical rules. Lewis calls such items "chunks" and differentiates between two main categories: collocations and expressions (Lewis, 1997). Collocations are words that co-occur with other words (e.g. "to brush one's teeth" rather than "to wash one's teeth" as in other languages). Concerning expressions, Lewis differentiates between fixed expressions such as "good morning" and semi-fixed expressions with one or more open slots that allow for variation, for example, "I would like...".

In our study, we regard both fixed and semi-fixed expressions as defined by Lewis as "chunks." We also consider one-word items such as "thanks!" to be fixed expressions. In the intervention concept, collocations do not form their own sub-category, but are instead introduced as parts of expressions.

Teachers who apply the Lexical Approach introduce chunks of differing length already at an early learning stage; the grammar is, however, not analyzed (Lewis, 1993). The chunks are taught as part of an action-oriented and needs-based approach, thereby quickly providing learners with the linguistic resources to communicate in daily life. It is expected that chunk-based teaching strategies promote fluency in language use because they make available larger units of language that can be recalled as a whole, a factor in successful communication. Compared to conventional teaching methods, learning and practicing sentence grammar and memorizing isolated words have little place in the Lexical Approach. Instead, the emphasis is placed on training listening and later reading comprehension. Learners are furthermore encouraged to engage in active learning using notebooks, in particular a personalized *lexical notebook* containing useful chunks (Lewis, 1997).

The Lexical Approach has received varying support from experts in the field of language teaching. Overall, the large amount of chunks in any language is widely accepted. For instance, Wray & Perkins (2000) point to the fact that up to 70% of the vocabulary of an adult native speaker (of English) consists of formulaic expressions, including collocations. There are also occasional references to how learning chunks can promote communicative language use already at an early stage in the acquisition process. If chunks are learned as lexis, they can soon be recalled as (correct) communicative resources, even if knowledge about linguistic forms is largely missing. If chunks lead to better communication, the number of learning opportunities may increase (Aguado,

2002; Nation & Newton, 2009; Wray, 2002). As concerns the acquisition of implicit and explicit language rules, it is expected that learners will gradually, or after a certain developmental phase, analyze their mental collection of formulaic expressions in terms of regularities (Lewis, 1997; Wray, 2000). It is furthermore posited that the results of this analysis will be stored in addition to the original chunk and also be available in language use (Wood, 2002). Whether and when this analysis takes place is a decisive question. Taking Usage-Based Linguistics as a basis, Bybee (2008) suggests complementing the Lexical Approach with other teaching elements that help learners develop specific morphosyntactic constructions.

In the teaching concept for the intervention group, the Lexical Approach is complemented by fluency training employing repetition to support acquisition. This decision is also shaped by Nation's concept of the "four strands" of good language teaching (Nation & Meara, 2002). According to this concept, 25% of class time should be reserved for activities in which already familiar lexical units are practiced without a specific linguistic focus, but under slight time pressure. This method is devised to make the learned units available for fluent recall. Two types of activities are differentiated: in the first, the familiar items are repeated several times (e.g. under increasing time pressure); in the second, enriching lexis is the goal: familiar vocabulary is experienced in new combinations and applications (e.g. through reading longer but simplified texts). Only the first type was applied in the intervention group.

3 Study Aims, Research Questions

The first goal of the study was to develop a workable course concept for the intervention group and its subsequent implementation by adapting the teaching materials. Furthermore, the study depended on the willingness of the language schools to work according to our instructions for two to four months. The main focus of the study was dedicated to the learning progress of the course participants. This progress was captured using suitable instruments, and then analyzed.

The study aimed to answer the following questions:

- i. Is it possible to develop an effective teaching concept, including the corresponding materials, in accordance with the basic principles of the intervention concept and to implement it in low-threshold courses for German as a second language?
- ii. Does the intervention concept meet with interest and general acceptance among teachers and learners?
- iii. Is the intervention concept suitable for learners without much previous schooling?
- iv. What learning outcomes are attained in different partial areas among learners in the intervention group as compared to the learners in the control group?

Interviews with the teachers were planned to answer questions 1 through 3.

Oral and written tests were given to determine the learning outcome. With regard to the results, the following expectations were posited: the intervention concept would be regarded as more successful than the control concept if the intervention group (IG) achieved the following results as compared to the control group (CG):

- iv.a. More frequent use of chunks in the target language
- iv.b. More fluent use of language
- iv.c. Somewhat higher accuracy in oral productions (feel-based accuracy); accuracy in written productions the same or only slightly lower (rule-based accuracy)
- iv.d. Comparable lexical knowledge (semantic level) and comparable skills in communicating content (not accounting for formal accuracy)

4 Methods

4.1 Study Design

The study was conducted as a quasi-experiment with an intervention and a control group in exclusively low-threshold⁷ intensive German courses. The participants attended the courses four to five days a week, for eight to eleven hours per week. Depending on the course provider, teachers were given between eight and sixteen weeks to work through the textbook.

To reduce the impact of these differences as well as of other potential, school-specific variations, a pair of parallel classes was formed each time a beginners'⁸ German course started. One of the classes was then taught according to the intervention concept, the other class according to the control concept. A stratified random sampling was used to assign the students to either the intervention or the control group. The schools' knowledge of the students was drawn on to best balance potentially relevant learner characteristics in both classes; these characteristics primarily concerned origin, native language, and educational background.

The schools then assigned the classes to teachers. Overall, there were seven pairs of classes at four schools. One pair later withdrew from the study because the criteria applied when assigning students to groups resulted in highly heterogeneous classes, causing dissatisfaction in the groups and necessitating a distribution according to level of language skills.

Our intervention took place in two phases of equal duration. The first phase could be considered a preliminary phase in which many learners made their first experiences with the German language and the teachers familiarized themselves with our concepts and materials. At the end of the first phase, an initial examination was given that was nearly identical with the final examination in terms of structure. After the first phase, some students left the group and others entered. All newly entering students took the initial examination. As a rule, the final examination was given in the last week of the second phase. The quantitative analysis therefore reflects the learning outcome of the second part of the course.

⁷ Access to low-threshold courses is made easier in various regards, especially regarding the degree of prior experience in learning and languages, and regarding cost.

⁸ This term is used in reference to the school's assessment of prospective students for an A1.1 course in accordance with the CEFR (Council of Europe & Goethe-Institut Inter Nationes etc., 2001).

4.2 Classroom Intervention

The intervention in the context of this study concerns many aspects of overall course design. The intervention had been thoroughly tested in earlier intensive courses and then refined and adapted based on the findings. In the intervention phase, our implementation measures included introductory workshops with school coordinators and teachers, class visits, discussions, teaching guidelines, teacher's journals, and, most importantly, the textbook materials, which prescribed many aspects and activities. Both the intervention group and the control group were affected by these measures, although the control group was impacted to a lesser degree.

For each pair of classes (IG and CG) the setting was comparable in a number of ways: the original textbooks were identical (therefore the topics remained the same); both groups had the same number of classroom hours and the same classroom infrastructure; also, the amount of homework was jointly determined by the teachers of the respective pair of classes. Depending on the concept applied, however, systematically different priorities were set and other teaching materials were sometimes used.

The rules of each concept were also binding for the teachers when they designed their own activities. To reduce mixing concepts as much as possible, the project team aimed to supply enough concept-appropriate material that was ready for use.

The two textbooks used were meticulously modified and subsequently reprinted for the intervention group. Certain elements were removed, changed, or complemented to ensure that working with chunks took on a dominant role. The following presents a summary of the changes to the teaching concept and the textbooks:

- Introducing daily fluency training sessions lasting 20 to 30 minutes. Such sessions repeat earlier content and the content of the previous day to equal parts. In fluency training, it is critical that learners be already familiar with the language items and that they have practiced them in class. In fluency tasks, it is often the case that activities from the textbook are simply repeated, although with the introduction of time pressure⁹ or at least the awareness that fluency is at stake. "Pointing" represents a versatile method to conduct a fluency

⁹ The trials had already demonstrated that time pressure in certain activities (e.g. dialogues) can cause stress and therefore result in negative attitudes. For this reason, no activities foregrounding time pressure were designed, such as the exercise Nation recommends, in which learners repeat the same monologue within an increasingly shorter period of time.

exercise: a keyword, a flashcard, or a picture serves as the basis for language productions of varying length and kind. Fluency training also includes written activities, for example, a group competition in which learners write as many sentences as possible about a picture.

- Replacing vocabulary lists as well as grammar tables with a *lexical notebook*, as Lewis (1997) recommends. Over the course of time, learners use this notebook to compile numerous, individualized expressions for their own communicative needs.
- Complementing the textbook with so-called everyday dialogues (only in the case of one textbook due to the amount of content available). Everyday dialogues were also made available to the control group but used in a different way. The turns and chunks occurring in the dialogues were treated more intensely in the intervention group. The everyday dialogues were also supplied to the learners as flash cards, which presented the dialogues schematically and with gaps, leaving room for personalization.
- Systematically adjusting the exercise activities. Whereas the control group worked with words or morphemes, the intervention group worked with chunks. Exercise types recommended by Lewis (1997) formed the basis:
 - Categorizing chunks – e.g. forming chunks that are relevant to the learners' communicative needs
 - Matching – e.g. matching “mixed up” parts of various dialogues
 - Completing – e.g. gap-fill exercises in which entire chunks are to be filled in, or in which useful fixed expressions can be learned. This type of exercise is extended to everyday dialogues in which entire roles must be completed.
- Replacing the numerous quizzes focusing on grammar with analogous instruments having a lexical (chunk) focus.

Certain activities typical of the original textbooks (i.e. also of the CG's instruction) were, for the most part, eliminated from the materials for the IG and replaced with other types of activity. This specifically concerns activities designed to introduce grammar rules, activities in which grammatical phenomena are to be identified and matched, gap-fill activities for morphology and syntax as well as exercises in which grammatical features are to be converted.

The CG materials were only slightly altered to include some additional material. The teachers were obliged to teach German in a way that adhered to the textbook's design and principles (as laid down in the concept for the CG) as closely as possible. As a consequence, teachers were not allowed to bring in their own materials and ideas as they normally would. As such, the CG's courses were more clearly defined and very similar to one another, despite having different teachers.

4.3 Data Collection

Three types of data collection were conducted: 1) gathering data on course participants based on information provided by the schools; 2) oral and written tests at two times (initial examination before and final examination after the second stage of the course); and 3) interviews with the teachers.

Data on course participants

All schools record the personal data of participants as well as diverse background information that is necessary for administrative purposes or for assigning students to a course. The linguistic-cultural and educational backgrounds of the participants were of particular interest in forming groups for our study. Other information on perceived familiarity with learning and ability to learn – based on the experience of former teachers – was only occasionally available and could therefore not be drawn on when forming groups or when analyzing the data (as variables explaining differing learning outcomes).

Tests

As a basic principle, the exams covered only knowledge and skills that the course participants could have acquired in the course. Potential influences from the learners' private lives were minimized where possible. A separate test was created for each textbook; the main part of the test items was, however, identical. The final examination consisted of items that had already been tested in the initial examination as well as of items related to content treated in the second part of the course. The test design using overlapping items allowed us to present all results per tested skill or competence using a single scale.

Trained test administrators conducted the examinations in the classrooms. The written test was a group examination that lasted approximately 80 minutes. The oral test was completed entirely on a computer. The test takers were individually assisted

by a test administrator. The first oral exam lasted 30 to 40 minutes, the second 35 to 45 minutes per course participant.

The following table provides an overview of the tested areas (test constructs) as well as of the testing instruments that were used in both tests:

10 The C-Test is a type of cloze test. Starting with the second sentence, the second half of every second word must be filled in (C-Test-Team, n.D.).

11 In our implementation of the elicited imitation tasks (Erlam, 2009) using a computer (with the software program *PsychoPy*, Peirce, 2009), the participants heard a single word or a sequence of words, then read a number on the computer screen. After hearing a signal, they had to repeat what they had heard. The number acts as a disruption, thus preventing the brain from regurgitating the input from its sensory memory (Shiffrin, 2003).

12 The speech rate and the other two measurements function as indicators for the fluency of spoken language (utterance fluency) (Segalowitz, 2010). Because speech rate combines both the articulation rate and the length of pauses, it represents a technical measurement of fluency that harmonizes well with perceived fluency. This is not the case with the articulation rate, which only measures speed (Kormos & Dénes, 2004; de Jong, Steinel, Florijn, Schoonen & Hulstijn, 2012).

13 No articulation for more than 250 milliseconds was deemed a silent pause.

Tested Area / Test Construct	Task Type [Coded Aspect]
Written Test	
1a) General written skills (focus on content)	C-Test ¹⁰ using 4 texts with 15 gaps [at least a recognizable reconstruction of the original word]
1b) General written skills (focus on content and form)	C-Test using 4 texts with 15 gaps [complete reconstruction of the original word]
2) Grammar: morphosyntax (potentially rule-based) Grammar: morphology (potentially rule-based) Lexicogrammar: prepositions (potentially rule-based)	Forming sentences from elements Varying sentences Naming the plural form Completing sentences
Oral Test	
3a) Comprehension and reproduction of spoken language (focus on content)	Elicited imitation ¹¹ [number of reproduced elements of content] Elicited imitation [number of reproduced word stems]
3b) Comprehension and reproduction of spoken language (focus on content and form)	Elicited imitation [number of completely reproduced words] Elicited imitation [number of reproduced chunks, (either complete or reduced to their core)]
4) Formulaic character of spoken utterances (chunking)	Elicited imitation [completeness and correctness of chunks]
5) Grammar: morphology (mainly feel-based)	Elicited imitation [morphosyntactical correctness]
6) Spoken fluency (cognitive fluency)	Guided oral production; guidance via images and recordings as linguistic models [a. speech rate ¹² – number of syllables per second in a selected stretch of language (including all pauses) ¹³ [b. pruned speech rate – number of syllables per second in a selected stretch of language (including all pauses) not counting syllables in filled pauses, repetitions, and self-corrections] [c. articulation rate – number of syllables per second, considering the sounding parts of the selected stretch only (no silent pauses)]

Table 1: Overview of the test constructs and the testing instruments used in the initial examination and the final examination

Interviews with teachers

The collaboration with the teachers was finalized with an interview after the experimental phase was concluded. The interviews with teachers from the intervention and the control groups were structured the same, with the exception of questions concerning the concept. Two main thematic areas were treated: 1) the concept and 2) the participants. The interviews lasted between 30 and 100 minutes.

All 16 interviews were transcribed and evaluated on the basis of a content analysis (Mayring, 2008) with reference to aspects of relevance to the study's main research questions.

- i. Implementation of the concept
 - a. Comparison with the teachers' regular methodology
 - b. Concrete implementation of the concept
 - c. Factors of influence
- ii. Work with the individual learners or the whole class
- iii. Advantages and disadvantages of the concept
 - a. Perspective of teachers
 - b. Perspective of students (as assessed by the teachers)

Processing test results

The written and oral test results were processed using teams of two coders, with reference to exact, written guidelines for data analysis.

Work on establishing the fluency indicators based on monologic speech sequences was particularly demanding. The audio recordings were segmented, categorized, and quantified with the help of the software program *Praat* (Boersma & Weenink, n.D.), but additional manual checks and corrections were necessary.

5 Results

5.1 Test Results

The results on the written examinations and on the elicited imitation test were Rasch-scaled.¹⁴ This guaranteed comparability of the test results gathered by means of four test forms (two textbooks; two times).¹⁵ The indicators of spoken fluency could be used directly, as they are based on the same oral production task.

The statistical analysis is based on repeated-measures (times 1 and 2) one-way (group membership) analyses of covariance (ANCOVA),¹⁶ implemented as a mixed-model.

When evaluating the various test sections, data from slightly different numbers of participants (N) were available. Of the 153 students who entered the second stage of the course, data on 141 participants were available from at least one test section, and data on 129 participants could be collected from all test sections. It should be noted that no significant correlation could be determined between test results on the initial examination and course drop-out.¹⁷ Table 2 provides an overview of the data; missing test data are mainly related to student absences due to organizational issues.

14 The statistical analyses were conducted using packages of the R software environment (Chambers, 2008; Bates, Mächler, Bolker & Walker, 2015; Kiefer, Robitzsch & Wu, 2015; Lenth, 2015; Halekoh & Højsgaard, 2014).

15 The results of all individuals were issued as so-called WLEs (Warm, 1989). WLE scales slightly overestimate the variability actually present in the data. To simplify the process, the WLEs were nevertheless treated as observed values.

16 In contrast to the analysis of variance, covariates were also factored in; it is known that such variables can influence individual differences. In our case, these factors concerned region of origin, duration of stay in Switzerland, education, and gender. Moreover, the variable "class pair" accounted for the fact that the schools simultaneously formed the intervention and control groups from a larger group of applicants who then learned German in a very similar environment.

17 This was tested using Pearson correlations between the variable for "Course drop-out – yes or no" and the test results on t1.

	IG	CG	Total
Course stage 2 begun	73	80	153
Course stage 2 ended (incl. at least 1 test section)	66	75	141
Women/men at the end of course stage 2	34/32	34/41	68/73
School education in years: mean (min.-max.)	8.8 (0-17)	9.3 (0-16)	9.1 (0-17)
Age: mean (min.-max.)	33.2 (16-51)	32.3 (17-57)	32.7 (16-57)
Candidates who took all test sections and partial tests on t2; data valid	63	66	129
Candidates who took written test section on t2	66	73	139
Candidates who took oral partial test I (elicited imitation) on t2	63	68	131
Candidates who took oral partial test II (oral production/fluency) on t2	64	71	135

Table 2: Overview of participants and test data

Table 3 provides an overview of the test results. For the three test sections, the results on a total of eleven areas tested are presented, including the mean values per group and per test stage. The standard error is given in brackets. Various significance tests control the generalizability of the results from a statistical point of view. The following is tested in each area under investigation: a) the improvement from t1 to t2 in IG and CG and b) the difference in improvement between the two groups. The data comparing the improvement given in the final, dark gray column is of central importance to our study.¹⁸

18 Two different p-values are given for the oral test sections. One (“p”) disregards that identical performance has repeatedly been coded and analyzed with reference to various phenomena. The other (“p_H”, “H” referring to a Holm-corrected value) observes the common, universal 95% confidence level per test section. As a result, Holm-corrected p-values are less frequently significant.

Investigated area	Intervention group (IG)			Control group (CG)			IG vs. CG
	Initial exam (t1)	Final exam (t2)	Significance of Improvement ¹⁹	Initial exam (t1)	Final exam (t2)	Significance of Improvement	Difference in Improvement; Significance ²⁰
Written test section N = 139, n (IG) = 66, n (CG) = 73							
1a) General written skills: written C-Test / complete reconstruction	0.064 (0.126)	0.434 (0.126)	t = 5.13 p < 0.0001	0.134 (0.113)	0.540 (0.113)	t = 5.92 p < 0.0001	F = 0.13 ²¹ p = 0.719
1b) General written language skills: written C-Test / correct word	0.105 (0.081)	0.497 (0.081)	t = 4.22 p = 0.0003	0.077 (0.074)	0.448 (0.074)	t = 4.19 p = 0.0003	F = 0.03 p = 0.864
2) Gramm. accuracy in written test (potentially rule-based): exercises testing morphology and syntax	0.017 (0.181)	0.469 (0.181)	t = 3.79 p = 0.0013	0.386 (0.163)	0.829 (0.163)	t = 3.91 p = 0.0008	F = 0.003 p = 0.953
Oral test section I: elicited imitation N = 131, n (IG) = 63, n (CG) = 68							
3a) Comprehension and reproduction of spoken language: number of reproduced content elements	-0.143 (0.131)	0.236 (0.131)	t = 5.74 p < 0.0001	-0.159 (0.119)	0.301 (0.119)	t = 7.24 p < 0.0001	F = 0.776 p = 0.380 p _H = nsig.
3b) Comprehension and reproduction of spoken language: number of completely reproduced words	0.055 (0.110)	0.379 (0.110)	t = 5.89 p < 0.0001	0.025 (0.099)	0.495 (0.099)	t = 8.88 p < 0.0001	F = 3.64 p = 0.059 p _H = nsig.
3c) Comprehension and reproduction of spoken language: number of reproduced word stems	-0.126 (0.135)	0.229 (0.135)	t = 5.42 p < 0.0001	-0.164 (0.122)	0.276 (0.122)	t = 6.99 p < 0.0001	F = 0.89 p = 0.348 p _H = nsig.
4) Formulaic oral expressions: chunking	-0.099 (0.119)	0.098 (0.119)	t = 3.27 p = 0.007	-0.160 (0.108)	0.205 (0.108)	t = 6.31 p < 0.0001	F = 4.075 p = 0.046 p _H = nsig.
5) Grammatical accuracy (rather feel-based): accuracy of morphosyntax	-0.029 (0.142)	0.248 (0.142)	t = 3.50 p = 0.0035	-0.196 (0.129)	0.455 (0.129)	t = 8.545 p < 0.0001	F = 11.566 p = 0.0009 p _H = 0.0045
Oral test section II: oral production N = 135, n (IG) = 64, n (CG) = 71							
6a) Spoken fluency: number of syllables per second in recording: speech rate	1.490 (0.054)	1.610 (0.054)	t = 2.73 p = 0.035	1.527 (0.049)	1.516 (0.049)	t = -0.29 p = 0.99	F = 4.765 p = 0.031 p _H = 0.093
6b) Spoken fluency: number of syllables per second in recording; excluding syllables in corrections, repetitions, and filled pauses; pruned speech rate	1.190 (0.053)	1.331 (0.053)	t = 3.27 p = 0.0075	1.214 (0.048)	1.230 (0.048)	t = 0.39 p = 0.98	F = 4.42 p = 0.037 p _H = nsig.
6c) Spoken fluency: number of syllables per second in recording excluding silent pauses: articulation rate	3.006 (0.057)	3.075 (0.057)	t = 1.35 p = 0.53	3.118 (0.051)	3.086 (0.051)	t = -0.65 p = 0.92	F = 2.02 p = 0.157 p _H = nsig.

Table 3: Overview of test results

When the Holm corrections (“ p_H ” values) are factored in, only one variable shows a difference between the improvement in the IG and the CG: the accuracy of morphosyntax in the oral section (elicited imitation) in area 5, where the improvement in the CG is larger than that in the IG. Without factoring in multiple testing (“ p ” values), the following comparisons of improvement become significant or borderline significant (in italics):

- In favor of CG: 3b) number of completely reproduced words and 4) chunking
- In favor of IG: 6a) speech rate and 6b) pruned speech rate

Advantages for the control group can be observed in three of five areas assessing performance in elicited imitation; the results in both groups are comparable in the two remaining areas. A comparison based on content is also interesting: the improvement in the CG is greater where accuracy and complete responses were foregrounded (areas 3b, 4 and, particularly 5). By contrast, the difference between the groups was negligible where formal aspects were of lesser importance.

Improvement between the initial examination and the final examination is almost always significant *within* the groups. Exceptions (the four light gray fields) can be observed in the fluency measures: in the CG, all three fluency measures remained the same between t1 and t2. In the IG only the articulation rate (measurement of speech rate excluding silent pauses, i.e. in the sounding parts only) showed no change. The mean values for the speech rate and the pruned speech rate, however, increased significantly by 8% and 12%.^{20 21}

19 p-values on t test upon improvement according to group using Tukey’s corrective for multiple testing.

20 p-value on F test with a correction for mixed models (Kenward-Roger approximation).

21 For clarity reasons, no degrees of freedom are given. Essentially, they can be derived from group size.

5.2 Results from Teacher Interviews

The following section reports important findings from the final interviews held with teachers. The organization mirrors that of section 4.3.

Teacher implementation of the concept

To ensure comparability of all classes in the IG and CG, concepts for both groups were developed, introduced, and documented in the form of guidelines. The implementation was furthermore underpinned with advisory sessions, in-class monitoring, and access to teacher journals on the process. The interviews lead to the conclusion that teaching was substantially influenced by the concept applied. Materials supplied by the teacher appear to have played only a minor role, as did activities that did not adhere to the concept. The innovative approach in the IG of beginning each day with 20 to 30 minutes of fluency training was not always easy to implement, especially at the start of the course.

Reception of concepts by participants and classes

The classes did not have a uniform opinion of the intervention concept. The more dynamic social forms in some oral (repetitive) exercises met with resistance among some students (for instance, older learners), while others welcomed them. One teacher believes the increased interaction was the reason the atmosphere in her class was particularly constructive.

Students without much previous schooling and weaker students appear to have a better opinion of the concept in the intervention group than do learners with more previous schooling or learners who were aiming to improve their accuracy. The former group of learners was relieved to do without grammar teaching, while the latter doubted the effectiveness of fluency training with little to no corrections. Learners with more previous schooling also requested formal grammar training and the corresponding learning materials. They did, however, view the regular repetition of known material during fluency training as a chance to learn.

The CG classes showed negative reactions mainly with regard to learning vocabulary based on lists of words, as prescribed in the textbooks.

Both groups voiced a clear, positive opinion of a CD with everyday dialogues, which was created especially for this study. The CD was provided as additional material in two variations designed for the different teaching concepts.

Teacher assessment of advantages and disadvantages of the concepts

Several teachers in both groups felt limited by the obligation to apply a fixed concept and use only prescribed materials. In particular, teachers who used the textbook that was less specifically adapted to the target audience would have preferred to skip certain elements regarding themes, vocabulary, and grammar, and would have chosen to replace them with their own materials. By contrast, the teachers in the IG classes clearly believed that the intervention concept could not have been implemented without the specially designed classroom materials.

The teachers in the IG voiced both approval and criticism of the various components in the intervention concept. The importance assigned to oral language use, conducted mainly in fluency training, was welcome, but the teachers had misgivings about the scarcity of written activities. They also noted that individualization during fluency training might be a problem: learners would frequently make errors and practice these errors, and teachers had little opportunity to intervene.

Overall, teachers gave a positive assessment of the activities with chunks that replaced activities focused on grammar. Nevertheless, they simultaneously addressed the lack of explicit grammar instruction: learners with more school experience expected to receive grammar lessons and corresponding exercises, expectations that were not met. The IG teachers also missed using grammatical meta-language as an efficient instrument in correcting errors and in explaining certain aspects of language.

Teachers generally approved of the individualized lexical notebook (LN). Nevertheless, they qualified their approval due to the amount of extra work required to correct them; it should be noted that this work is essential because the chunks in the LN serve as the basis for the repetitions in fluency training. Teachers also pointed out that learners without much previous schooling required (too) much assistance in working with the LN. On a positive note, they pointed out that individualized elements such as the LN or the flashcards with everyday dialogues promoted learner autonomy.

In direct contrast to the concept in the intervention group, teachers believed the concept used in the control group (i.e. an approach based on the suggestions in the teaching manual) was geared toward fast learners and those with more previous schooling. Assessment of this condition is, however, not uniform. On the one hand, teachers believe the focus on grammar and grammatical terminology is too strong and there is too much explicit vocabulary work required. On the other hand, they claim this intensive vocabulary work is important. One teacher stresses that the CG concept should be retained to familiarize students without much previous schooling with the peculiarities of school learning.

6 Discussion

Overall, the findings yield a great deal of information on implementing the teaching concepts in class as well as on the linguistic development of the students in the course.

Teacher statements from the guided interviews present numerous possibilities to answer the first three research questions (cf. section 3). The new teaching concept and the related materials (question 1) were successfully developed and implemented. To be sure, some teachers in both groups found fault with the prescriptive methodology during the intervention phase and would have welcomed the opportunity to use their own materials that would directly address the needs of their students. From a research perspective, however, the prescriptive methodology was required to ascertain the influence the two teaching concepts had on the results.

The general acceptance accorded the intervention concept by teachers and learners (question 2) is less clear-cut and calls for a nuanced analysis. The concept excludes common elements of foreign language instruction (primarily explicit instruction in grammar) while emphasizing other factors (mainly daily fluency training as well as activities in which chunks are practiced in their functional contexts but without explicit analysis); as such, it does not meet expectations of students with experience in traditional foreign language learning or of students aiming to improve accuracy in their language skills. Moreover, the intervention concept presents a challenge to teachers, as it is at variance with past practices. One problem for teachers and some learners was posed by the fact that the fluency training automatized language that learners had only imperfectly mastered. The idea of fluency training is to only use language that is known and mainly correct (Nation & Meara, 2002). The students, however, could not rely on enough correct language, leading to the production of numerous errors.

The lack of explicit grammar lessons was identified as a further problem. Some students with a stronger educational background requested grammar lessons and/or exercises, while teachers said the lack of meta-language prevented them from giving appropriate feedback and making corrections. By contrast, weaker students and those without much previous schooling evidently benefited from the lack of explicit grammar instruction, as this reduced the learning load. The teachers' remarks on this aspect therefore permit an overall positive answer to the third research question on the suitability of this concept for students without much previous schooling.

As a rule, the heterogeneity of the learners poses a major challenge in teaching low-threshold language classes. The differing attitudes toward the lack of grammar instruction present the clearest evidence in this area. A possible consequence could be to use the intervention concept more frequently for learners without much previous schooling; likewise, the concept could be complemented by selected elements from conventional

teaching practices, especially explicit grammar instruction, which provides the terminology for corrections and enables students to work independently with other grammar resources. By the same token, fluency training done in accordance with Nation & Meara (2002) could be integrated into traditional teaching methods; this is of particular value because a stronger focus on correct language production reduces the danger of practicing incorrect language but also because the targeted development of fluency rarely features in conventional teaching concepts, despite the general recognition of fluency as an important aspect to develop and a core feature of language competence.

The results of the language tests correspond partially to the (optimistic) forecast made on the basis of research on the impact of the Lexical Approach (c.f. section 3 above).

The expectations were confirmed in three areas:

- Improvement in fluency is greater in the IG.
- No differences can be discerned in the improvement of vocabulary (semantic level) or in the ability to communicate content.
- There are no differences in the improvement in written accuracy.

The expectations were not confirmed in two other areas:

- The IG demonstrates less improvement in the application of correct and complete chunks than the CG.
- The CG demonstrates greater improvement in grammatical accuracy in oral language use.

It is of particular interest to formulate an overall interpretation of the results with reference to the interviews with the teachers.

The degree to which the overall findings are related to the teachers' statements on the use of incorrect language in fluency training (i.e. frequent activities of a highly repetitive nature) is unclear. Does the repetition of incorrect language lead to the memorization and production of incorrect chunks or constructions? The study's findings provide at least no clear counter-argument to this understanding, as the oral productions in the IG are less correct than those in the CG. This result is derived from comparisons in the three areas coded in elicited imitation output measuring aspects of accuracy: grammatical accuracy, precision of vocabulary, and correct chunking. At the same

time, the oral fluency of learners in the IG improved, in contrast to the CG. Significantly, this improvement is not due to faster articulation of the sounds of the language (articulation rate) but due to less pausing (i.e. pauses, fillers, repetitions, and self-correction) in relation to total speaking time. This condition is reflected in the results in the learner's speech rate and in their pruned speech rate. Longer pauses (or less productive speaking time) indicate that learners in the CG required more planning. The students in the IG, however, appeared to be relying on existing speech patterns, which indicates a prior proceduralization of the learned language (de Jong & Perfetti, 2011) or the availability of constructions that can be recalled. The results on written accuracy, which show no difference between the IG and CG, do not contradict this interpretation. It is plausible that perceived time constraints in the elicited imitation section allowed the "ballistic" nature of automatized language (Segalowitz & Hulstijn, 2009) to become manifest, while the written examination gave the students enough time to apply other knowledge with a corrective effect. An open question remains concerning the forms of knowledge applied: were the variously acquired chunks compared or did learners apply abstract (rule-based) knowledge derived from their experiences with language?

The results in the oral sections should also be viewed from the perspective of research findings on performance in speaking tasks, with particular reference to the Trade-Off Hypothesis (Skehan, 2009). The basic tenet of this hypothesis is the generally recognized condition that our cognitive capacity is restricted; specifically, the capacity of our working memory and our attentional capacity are both limited. As a result, competition arises between complexity, accuracy, and fluency²² in spoken language production.

In our highly structured test activities, the complexity factor was probably less important; it was also essentially the same for both groups. The cognitive resources could therefore be largely devoted to accuracy and fluency of language production. On this basis, it appears that learners in the IG focused on fluency while those in the CG favored accuracy. Yet proponents of the Lexical Approach (Nation & Newton, 2009) do not believe this outcome is inevitable. Instead, it is posited that acquired chunks will lead to both increased fluency and increased accuracy. Similarly, it is expected that, in cases with chunks requiring grammatical modifications, the application of known chunks – which are quickly and easily recalled – would free up cognitive resources, thus allowing the focus to be placed on modifications and corrections. Neither expectation was underpinned by our data. In the present case, it is conceivable that it is not the limited cognitive

²² The relationship between *complexity*, *accuracy*, and *fluency* (CAF) is the topic of an ongoing debate in applied linguistics and in research on language testing.

resources that are at stake, but rather the different characteristics of the language competence developed in the two groups. The IG probably had more automatized – although incorrect – chunks at their disposal, which their memories produced with “ballistic” force; by contrast, the learners in the CG possessed the necessary linguistic knowledge and awareness to produce somewhat more accurate German, but at the cost of fluency.

Our findings, as do many others, also contradict the non-interface position (Krashen, 1982): the acquisition of grammar for spontaneous application appears to succeed precisely *through* explicit instruction. It must be stated, however, that the intervention concept applied in the IG, in which grammatical structures are only present in frequently used chunks, yielded comparable improvements in form-focused written language use. Incidental learning apparently occurred, making available grammatical resources that could be applied in the written grammar items.

But what about the conjecture that chunk-based learning would, at a later period in time, allow learners to perform a grammatical analysis of already learned language? This is in principle conceivable. It depends on the significance of the observed difference between the IG’s oral and written accuracy from a cognitive point of view. Is the, relatively speaking, larger success in written language a sign that learners would have more accurate forms in store, but that they cannot (yet) automatically recall them, i.e. that these forms have not yet attained enough ballistic force? If so, it is entirely conceivable that successful analysis of linguistic forms could occur in future instruction, especially if that instruction emphasizes the production of formally accurate language.

Overall, it is significant that only learners in the IG demonstrated an improvement in fluency, as the data on both speech rates indicate. This confirms Lewis’s expectations (Lewis, 1993) and those of other proponents of the Lexical Approach. Fluency is a key feature in the development of language proficiency. The learners in the CG appear, however, to have reached an impasse in this area, although it is unclear whether this is solely a symptom of competition between fluency and accuracy, as conditioned by the brain’s limited memory capacity. This could very well be the case, as the results in the pruned speech rate, which reflects planning and self-correction, indicate that learners in the CG actually worked harder than those in the IG when producing language. In the future, the learning objectives in this group should include the development of automatized language (in the form of easily available chunks or constructions) through repetition activities. It is certain that language contact is easier when individuals have greater fluency; this in turn gives rise to more opportunities for learners to practice and acquire knowledge outside the classroom setting.

7 Summary

The findings of this study call for a nuanced response to the practical question underlying it. A chunk-based approach to teaching that does away with explicit grammar instruction (*focus on forms* and *focus on form*,²³ Long, 1991) can be successfully implemented in low-threshold language courses, but the approach meets with resistance among students with a stronger educational background and those looking to improve accuracy in their language production, while teachers claim students without much previous schooling tend to welcome it. Teachers themselves point to an important prerequisite to successfully implementing the intervention concept: appropriate course materials must be available. In addition, the same teachers stress a problematic issue: the frequent repetition of non-target-like language, which appears to be an unavoidable aspect of the intervention concept. It is precisely this aspect that most likely accounts for the weaker accuracy results in oral language production among learners in the IG in comparison to those in the CG, who learned using conventional methodology and numerous activities focusing on forms. But the learners in the CG made no significant advances in the equally key area of spoken fluency during the intervention phase, as did learners in the IG. These observations speak in favor of balanced learning opportunities (Ellis, 2005b). The findings indicate that the teaching concept in the IG requires at least a stronger emphasis on the accuracy of the chunks that are to be automatized. This does not necessarily need to be in the form of systematic grammar teaching, but would imply an overall stronger focus on forms – possibly simply greater insistence on correct forms in the target language that will later be practiced in fluency training. Conversely, the teaching concept in the CG lacks repetition exercises that target the promotion of automatization processes.

As concerns further research, an in-depth analysis of the same spoken productions that were used to determine the fluency indicators for the present study would be desirable in order to shed light on the differences between the two groups in the way they produce language and, in turn, on some of the cognitive processes involved in their speech production. For example, do the groups demonstrate differences in the kind of formal errors produced? Can differences be detected regarding the kinds of filled pauses, of corrections, and repetitions? It would also be of value to investigate the impact fluency training has on regular foreign-language learning at school, where courses are generally less intensive and where fewer opportunities to practice language exist outside the classroom. In such a study, differing levels of skills should be consid-

23 *Focus on forms*: systematic explicit grammar instruction; *focus on form*: explicit grammar instruction based on needs arising in a communicative language teaching setting.

ered, thus making allowances for the condition that the size of an individual's repertoire of correct forms before fluency training starts may lead to differing results in the development of formal accuracy.

8 Bibliographie Bibliografia Bibliography

Aguado, K. (2002). Formelhafte Sequenzen und ihre Funktionen für den L2-Erwerb. *Zeitschrift für Angewandte Linguistik (ZfAL)*, 37, 27–49.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C. & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111(4), 1036–1060.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48.

Boersma, P. & Weenink, D. (o. J.). *Praat. Doing phonetics by computer*. Abgerufen 3. Dezember 2013 von <http://www.fon.hum.uva.nl/praat/>

Bybee, J. (2008). Usage-based grammar and second language acquisition. In P. Robinson & N. C. Ellis (Hrsg.), *Handbook of cognitive linguistics and second language acquisition*. New York: Routledge/Taylor and Francis Group, 216–236.

Chambers, J. (2008). *Software for data analysis programming with R*. New York: Springer.

C-Test-Team. (o. J.). C-Test [Der Sprachtest]. Abgerufen 11. Oktober 2015 von <http://www.c-test.de/deutsch/index.php>

de Jong, N. H., Steinel, M. P., Florijn, A., Schoonen, R. & Hulstijn, J. H. (2012). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics*, 34(5), 893–916.

de Jong, N. & Perfetti, C. A. (2011). Fluency training in the ESL classroom: an experimental study of fluency development and proceduralization. *Language Learning*, 61(2), 533–568.

Ellis, N. C. (2005a). At the interface: dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27(2), 305–352.

Ellis, N. C. (2005b). Constructions, chunking, and connectionism: the emergence of second language structure. In C. J. Doughty & M. H. Long (Hrsg.), *The handbook of second language acquisition* (New edition). Malden, MA: Wiley-Blackwell, 63–103.

Erlam, R. (2009). The elicited oral imitation test as a measure of implicit knowledge. In R. Ellis, S. Loewen & C. Elder (Hrsg.), *Implicit and explicit knowledge in second language learning, testing and teaching*. Multilingual Matters, 65–93.

Europarat & Goethe-Institut Inter Nationes etc. (Hrsg.). (2001). *Gemeinsamer europäischer Referenzrahmen für Sprachen: lernen, lehren, beurteilen*. Berlin; Zürich etc.: Langenscheidt.

Funk, H., Kuhn, C., Demme, S. & Bayerlein, O. (2005). *Studio d A1. Schweizer Ausgabe. Kurs- und Übungsbuch Deutsch als Fremdsprache*. (1. Auflage, 2. Druck). Berlin: Cornelsen Verlag.

Gatbonton, E. & Segalowitz, N. (2005). Rethinking communicative language teaching: a focus on access to fluency. *Canadian Modern Language Review/ La Revue canadienne des langues vivantes*, 61(3), 325–353.

Halekoh, U. & Højsgaard, S. (2014). A Kenward-Roger approximation and parametric bootstrap methods for tests in linear mixed models – the R package pbkrtest. *Journal of Statistical Software*, 59(9), 1–30.

Kiefer, T., Robitzsch, A. & Wu, M. (2015). *TAM: Test Analysis Modules*. Abgerufen 10. Mai 2015 von <http://cran.r-project.org/package=tam>

Kormos, J. & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32(2), 145–164.

Krashen, S. D. (1982). *Principles and practice in second language acquisition* (1st ed.). Oxford; New York: Pergamon.

Lenth, R. (2015). *Ismeans: least-squares means*. Abgerufen 10. Mai 2015 von <http://cran.r-project.org/package=lsmmeans>

Lenz, P., Andrey, S. & Lindt-Bangerter, B. (2009). *Rahmencurriculum für die sprachliche Förderung von Migrantinnen und Migranten*. Bern: Bundesamt für Migration BFM. Abgerufen 7. März 2013 von <http://www.fide-info.ch/de/fide/rahmencurriculum>

Lewis, M. (1993). *The lexical approach: the state of ELT and a way forward*. Hove: Language Teaching Publications.

Lewis, M. (1997). *Implementing the lexical approach: putting theory into practice*. Hove: Language Teaching Publications.

Lewis, M. (Hrsg.). (2000). *Teaching collocation: further developments in the lexical approach*. Hove: Language Teaching Publications.

Long, M. H. (1991). Focus on form: a design feature in language teaching methodology. In K. de Bot, R. B. Ginsberg & C. Kramsch (Hrsg.), *Foreign language research in cross-cultural perspective* (Bd. 2). Amsterdam: John Benjamins Publishing Company, 39–52.

- Long, M. H. (2005). Stabilization and fossilization in interlanguage development. In C. J. Doughty & M. H. Long (Hrsg.), *The handbook of second language acquisition* (New edition). Malden, MA: Wiley-Blackwell, 487–535.
- Mayring, P. (2008). *Qualitative Inhaltsanalyse. Grundlagen und Techniken* (10.). Weinheim und Basel: Beltz Verlag.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge etc.: Cambridge University Press.
- Nation, P. & Meara, P. (2002). Vocabulary. In N. Schmitt (Hrsg.), *An introduction to applied linguistics*. London: Hodder Arnold Press, 35–54.
- Nation, P., & Newton, J. (2009). *Teaching ESL/EFL: listening and speaking*. New York: Routledge/Taylor and Francis Group.
- Nattinger, J. R. & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford; New York: Oxford University Press.
- Niebisch, D., Penning-Hiemstra, S., Specht, F. & Bovermann, M. (2011). *Schritte plus 1. Niveau A1/1. Ausgabe Schweiz. Kursbuch + Arbeitsbuch*. (1. Auflage). München: Hueber Verlag.
- Norris, J. M. & Ortega, L. (2001). Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning*, 51, 157–213.
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2(10).
- Segalowitz, N. (2005). Automaticity and second languages. In C. J. Doughty & M. H. Long (Hrsg.), *The handbook of second language acquisition* (New edition). Malden, MA: Wiley-Blackwell, 382–408.
- Segalowitz, N. (2010). *Cognitive bases of second language fluency*. New York: Routledge.
- Segalowitz, N. & Hulstijn, J. (2009). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. de Groot (Hrsg.), *Handbook of bilingualism: psycholinguistic approaches* (1. Aufl.). Oxford: Oxford University Press, 371–387.
- Shiffrin, R. (2003). Modeling memory and perception. *Cognitive Science*, 27(3), 341–378.
- Skehan, P. (2009). Modelling second language performance: integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510–532.
- Spada, N. & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: a meta-analysis. *Language Learning*, 60(2), 263–308.
- Warm, T. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54(3), 427–450.

- Willis, D. (1990). *The lexical syllabus: a new approach to language teaching*. London; Glasgow: Collins ELT.
- Wood, D. (2002). Formulaic language in acquisition and production: implications for teaching. *TESL Canada Journal*, 20(1), 1–15.
- Wray, A. (2000). Formulaic sequences in second language teaching: principle and practice. *Applied Linguistics*, 21(4), 463–489.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge; New York etc.: Cambridge Univ. Press.
- Wray, A. & Perkins, M. R. (2000). The functions of formulaic language: an integrated model. *Language & Communication*, 20(1), 1–28.

