

Zwei (Aus)Sprachen? Die Entwicklung eines neuen Korpusprojekts für die bilinguale Phonologie des Spanischen und Katalanischen

1. Die Rolle von Korpora in der empirischen Sprachwissenschaft

Korpora sind nach bestimmten Kriterien zusammengestellte Sammlungen von sprachlichen (mündlichen oder schriftlichen) Äußerungen, die als empirische Grundlage für sprachwissenschaftliche Untersuchungen dienen können. Auch wenn zahlreiche andere Möglichkeiten zur Datengewinnung existieren (Befragungen, Experimente etc.), spielen Korpora noch immer eine zentrale Rolle in der empirischen Sprachwissenschaft. Gut aufbereitete Korpora, die in digitaler Form *open-access* verfügbar sind, sind aus diesem Grund besonders wertvoll für die linguistische Forschung.

Das Ziel des hier vorgestellten Projekts ist es, ein neues Korpus zu erstellen, das der Untersuchung der Aussprache von bilingualen SprecherInnen des Spanischen und Katalanischen ist. Innovativ ist dabei, dass es als erstes Korpus vergleichbare Daten von denselben bilingualen SprecherInnen in beiden Sprachen enthält. Zudem soll es unter *open-access* Bedingungen öffentlich zugänglich gemacht werden.

2. Spanisch-Katalanischer Sprachkontakt

Das Spanische und Katalanische stehen in einem engen Kontaktverhältnis. Fast im gesamten katalanischsprachigen Raum werden sowohl Spanisch als auch Katalanisch gesprochen, weshalb sich beide Sprachen gegenseitig beeinflussen. Gleichzeitig existieren in beiden Sprachen interne Wandelprozesse, die unabhängig voneinander stattfinden. Diese Situation der gegenseitigen Beeinflussung und internen Wandels stellt eine Herausforderung für linguistische Forschung dar, der die aktuelle Korpusphonologie noch nicht zufriedenstellende begegnen kann.

3. Stand der Forschung

Sowohl für das Spanische als auch für das Katalanische existieren bereits zahlreiche phonologische Korpora sowie Korpora, die speziell für die Untersuchung des Spanisch-Katalanischen Sprachkontakts entwickelt worden sind (Sinner 2001, Vann 1995, Lleó 2011). Jedoch ist all diesen Korpora gemein, dass sie entweder Daten nur auf Spanisch oder auf Katalanisch enthalten oder Daten aus beiden Sprachen, aber von unterschiedlichen SprecherInnen zusammenfassen. Bislang fehlt ein bilinguales Korpus, d.h. ein Korpus, das Sprachaufnahmen in beiden Sprachen von denselben bilingualen SprecherInnen in beiden Sprachen enthält. Nur ein solches Korpus ermöglicht es, mögliche Transferphänomene der einen in die andere Sprache festzustellen und die zugrundeliegenden Phänomene in der L1 (der Erstsprache) zu lokalisieren. Das hier vorgestellte Projekt beabsichtigt, dieses Desideratum einzulösen.

4. Entwicklung eines neuen Korpusprojekts

Ausgangspunkt für das neue Korpusprojekt ist bereits existierende Korpus IFEC (*Inter-)*Fonología del Español Contemporáneo IFEC (Pustka/Gabriel/Meisenburg/Burkard/Dziallas 2018), das die

Dokumentation der Aussprache des Spanischen in der Welt zum Ziel hat. Das zugehörige Protokoll enthält zum einen einen soziodemographischen Fragebogen, das neben Fragen zu Alter, Beruf etc. auch Fragen zur sprachlichen Biographie enthält. Die sprachlichen Daten werden durch das Lesen einer Wortliste, eines Textes und durch ein semistrukturiertes Interview erhoben. Dieser Mix aus elizierter und (Semi-)Spontansprache verspricht ein hohes Maß an Komparabilität und Authentizität.

In einem ersten Schritt wurde die spanische Wortliste um Wörter erweitert, die häufig dokumentierte Transferphänomene enthalten können. Der zweite Schritt bestand aus der Erstellung eines analogen Korpus für das Katalanische. Dafür wurden zunächst die wichtigsten phonologischen Phänomene des Katalanischen und dessen intradialektalen Varietäten systematisiert. Anschließend wurden zunächst für diese Phänomene Beispielwörter gesucht, die die entsprechenden lautlichen Kontexte enthalten und dann ein Text verfasst, das ein Viertel der Wörter aus der Wortliste enthält. Wie auch das Protokoll für das Spanische enthält auch das neue Protokoll ein semistrukturiertes Interview, wofür die Fragen aus dem Spanischen ins Katalanische übersetzt wurden.

Am Ende wurde das gesamte Protokoll von Muttersprachlern aus verschiedenen Regionen Korrektur gelesen, um sicherzustellen, dass der Text und die Interviewfragen dialektübergreifend verständlich ist.

5. Erste Datenerhebung auf Barcelona und Menorca

Dieses neu erstellte Protokoll wurde (zusammen mit dem bereits existierenden IFEC-Protokoll) im Jahr 2016 bereits in zwei Untersuchungsorten (Barcelona und Menorca) angewandt. Beide Regionen sind zweisprachig, jedoch ist auf Menorca das Katalanische weitaus präsenter als in Barcelona. So bezeichnen nach Umfragen aus den Jahren 2013 bzw. 2014 in Barcelona 34,6 % das Katalanische als ihre Gewohnheitssprache *lengua habitual* (Idescat 2013), während auf Menorca der Prozentsatz bei 54 % liegt (EULIB 2014).

Entsprechend den Vorgaben des FEC-Protokolls wurden in beiden Orten jeweils 12 bilinguale SprecherInnen aufgenommen, die drei verschiedenen Altersgruppen angehören.

6. Ein Anwendungsbeispiel: das velarisierte /l/ im Spanischen und Katalanischen

In diesem Abschnitt soll anhand eines konkreten Phänomens ein Anwendungsbeispiel des Korpus skizziert werden. Während das /l/ im Spanischen gewöhnlich wie im Deutschen gebildet wird (in dem die Zungenspitze zum Zahndamm *Alveolen* bewegt wird), wird bei der Artikulation des katalanischen /l/ zudem der Zungenrücken an das Gaumensegel angehoben. Man spricht bei dem resultierenden Laut ([ɫ]) von einem *velarisierten* /l/, wobei der Grad der Velarisierung von der jeweiligen lautlichen Umgebung abhängt (Recasens 1986).

Diversen Studien zufolge lassen sich jedoch sowohl im Spanischen, als auch im Katalanischen Abweichungen von dieser Norm feststellen. Nach Wesch 1997 existiert auch im Spanischen Barcelonas die Variante eines velarisierten /l/, das vor allem von älteren SprecherInnen produziert wird. Gleichzeitig beobachtet Julià i Muné 2008 die Herausbildung eines Sprachstils, der vor allem in höheren sozialen Klassen verwendet wird und der durch ein 'klares', nicht-velarisierendes /l/ charakterisiert ist.

Um zu überprüfen, ob in diesen Fällen tatsächlich Transfer vorliegt oder ob sie durch interne Wandelprozesse erklärt werden können, müssen für die Analyse empirische Daten beider Sprachen herangezogen werden. Hierzu wurden zunächst die Sprachaufnahmen der Wortlisten aus Barcelona und Menorca bezüglich Frequenz, Velarisierungsgrad (der mithilfe der Software PRAAT gemessen wurde) ausgewertet.

Die Ergebnisse der Katalanischen Wortlisten zeigen, dass das Phänomen der /l/-Velarisierung im Katalanischen durchaus noch sehr präsent ist, sowohl in Barcelona, als auch auf Menorca. Jedoch lässt sich in beiden Orten ein Abwärtstrend beobachten: je jünger die SprecherInnen sind, desto weniger /l/-Velarisierungen sind beobachtbar, dabei ist der Velarisierungsgrad auf Menorca allgemein höher als in Barcelona. In den spanischen Wortlisten zeigt sich ein ähnlicher Abwärtstrend wie in den katalanischen Daten. Jedoch gibt es hier große Unterschiede zwischen den beiden Untersuchungsorten: in Barcelona ist der Velarisierungsgrad zum einen wesentlich geringer als in den katalanischen Daten, zum anderen lassen sich unter den jüngeren SprecherInnen kaum Velarisierungen beobachten (d.h. vor allem jüngere SprecherInnen scheinen Aussprache zu trennen). Auf Menorca ist der Velarisierungsgrad dagegen in den spanischen Daten ähnlich hoch wie in den katalanischen Daten, was sich womöglich darauf zurückführen lässt, dass es auf Menorca zu weniger Kontakt zu monolingualen SpanischsprecherInnen kommt.

Die Vorteile des bilingualen Korpus werden hier besonders deutlich, wenn man die Daten aus Barcelona betrachtet. Bei SprecherInnen mit L1 Katalanisch wäre es durchaus erwartbar, dass sie die /l/-Velarisierung ins Spanische transferieren. Unsere Ergebnisse zeigen, dass dies nur bei älteren SprecherInnen der Fall ist. Der Vergleich mit den katalanischen Daten (der durch unser Korpus ermöglicht wird) zeigt, dass einige jüngere SprecherInnen bereits in ihrer L1 (in dem Fall Katalanisch) nicht-velarisierte /l/s produzieren. Statt zu untersuchen, warum ein Transfer eines Merkmals nicht stattfindet, müsste also hier untersucht werden, warum das jeweilige Merkmal bereits in der L1 nicht vorhanden ist.

7. Ausblick

Der nächste Schritt in der Korpusentwicklung ist die Ausdehnung auf weitere Untersuchungsorte (Tortosa, Lleida und Alicante). Nach der vollständigen Transkription aller Sprachaufnahmen sollen sämtliche Daten (Fragebögen, Aufnahmen inkl. Metadaten und derer Transkription) werden veröffentlicht. Zusätzlich soll auch das Protokoll öffentlich zugänglich bereitgestellt werden, sodass auf andere ForscherInnen es für die Datenerhebung in Orten verwenden können, die durch unser Korpus noch nicht erfasst sind.

Literatur

- Institut d'Estadística de Catalunya (IEC) (2015): *Estadística d'usos lingüístics a Catalunya 2013*. Idescat: Generalitat de Catalunya.
- Julià i Muné, Joan. (2008) Els son del català. In Joan Solà, Maria-Rosa Lloret, Joan Mascaró, & Manuel Pérez Saldanya (eds.), *Gramàtica del català contemporani*, Volum 1. pp. 37-87. Barcelona: Editorial Empúries.
- Lleó, Conxita/Cortés, Susana/Benet, Ariadna (2008): "Contact-induced phonological changes in the Catalan spoken in Barcelona", en: Siemund, Peter/Kintana, Noemi (eds): *Language Contact and Contact Languages*, Amsterdam: John Benjamins, 185-212.

- Pustka, Elissa/ Gabriel, Christop. y Meisenburg, Trudel (2016): Romance Corpus Phonology: from (Inter) Phonologie du Français Contemporain (I)PFC to (Inter-)Fonología del Español Contemporáneo (I)FEC. Póster presentado al P&P12 (13./14.10.2016), München, Alemania.
- Recasens, Daniel (1986): *Estudis de fonètica experimental del català oriental central*. Barcelona: Publicacions de L'Abadia de Montserrat.
- Sinner, Carsten (2001): Corpus oral de profesionales de la lengua castellana en Barcelona. <http://www.carstensinner.de/castellano/corpusorales/index.html>.
- Vann, Robert (2015): "Vann corpus 1995" DARDOSIPCAT <http://scholarworks.wmich/dardosipcat/transcriptions/>.
- Wesch, Andreas (1997): "El español hablado de Barcelona y el influjo del catalán - esbozo de un programa de investigación", in: *Verba Anuario Galego de Filoloxía* 24, 287-312.