

Vorabdruck von: Siebenhaar, Beat (2014): "Instrumentalphonetische Analysen zur Ausgestaltung des Sprechlagenspektrums in Leipzig". Zeitschrift für Dialektologie und Linguistik 81: 151–190.

BEAT SIEBENHAAR

INSTRUMENTALPHONETISCHE ANALYSEN ZUR AUSGESTALTUNG DES SPRECHLAGENSPEKTRUMS IN LEIPZIG¹

1 Forschungsüberblick

Das Obersächsische, in das die Mundart von Leipzig eingeordnet wird, nimmt in der Dialektologie des Deutschen eine Sonderstellung ein. Es ist der Dialekt, der am wenigsten Unterschiede zum Standard aufweist (FRINGS 1956, LAMELI 2013: 234–245) und wurde deshalb eine Zeit lang als Vorstufe zum Hochdeutsch gesehen. Dadurch, dass die ostmitteldeutschen Gebiete erst spät deutsch besiedelt wurden, sind sich in früher Siedlungszeit Sprecher unterschiedlicher Mundarten in diesem Raum begegnet. Aus dieser Mischung der Bevölkerung habe sich dann eine neue dialektale Ausgleichssprache ergeben. Im Rahmen der Kulturräumforschung zeichnet FRINGS (1932; 1936; 1956) diese Siedlungsgeschichte Sachsens anhand sprachlicher Zeugnisse nach. Seine Analyse der Daten aus dem Wenkerschen Sprachatlas ergibt durch mehr oder weniger deutlich übereinstimmende lautliche Merkmale eine dreiteilige Gliederung Sachsens. Diese lässt sich mit einer Besiedlung aus dem Westmitteldeutschen, aus dem Fränkischen und Nordbairischen und aus dem Niederdeutschen erklären. Dieser Befund FRINGS' gilt bis heute. Seine

¹ Der Autor dankt besonders Anja Schaufuß und Matthias Hahn für die Hilfe bei der Datenaufbereitung sowie den anonymen Gutachtern für die anregenden Vorschläge, die in den Beitrag eingeflossen sind.

weiterführende Interpretation, dass aus der Mischung dieser Mundarten in einer kolonialen Ausgleichssprache der direkte Vorläufer der deutschen Schriftsprache entstanden sei, wird heute aber klar zurückgewiesen (BESCH 2003). Ein Einfluss besonders städtischer Ausgleichsformen ist aber nicht von der Hand zu weisen (z. B. LERCHNER 2003, 2752–2755).

Von den 1930er bis in die 1960er Jahre sind einige Gebietsmonographien erschienen, die mit direkter Erhebungsmethode die im Wenker-Atlas dargestellte Variation der Grundmundarten präziser zu erfassen versuchen (BARTHEL 1933, BECKER 1933, BELLMANN 1962, FLEISCHER 1961, GROSSE 1955, PROTZE 1957). Die dort abgebildeten Karten zeigen häufig für ein lautliches Phänomen je nach Belegwort stark variierende Isoglossenverläufe. Die Autoren weisen deshalb immer wieder darauf hin, dass die ortsgenaue Ziehung der Isoglossen auch im ländlichen Raum im zentralen Gebiet Sachsens kaum mehr möglich sei, „weil die meisten Formen nicht mehr im täglichen Gebrauch erscheinen und deshalb einzig das sichere Urteilsvermögen, das Sprachrichtigkeitsgefühl, stark beeinträchtigt“ (GROSSE 1955: XV) sei bzw. sich sogar „aufgelöst“ (FLEISCHER 1961: 154) habe. Ein Basisdialekt war also schon in der Mitte des 20. Jahrhunderts kaum mehr wirklich zu erheben. Die Überblicksdarstellung von BECKER / BERGMANN (1969) hat dementsprechend im Bezug auf die Dialekte für das Gebiet im Dreieck Leipzig, Dresden, Chemnitz schon zur Publikationszeit weitgehend historischen Charakter, genau so wie die archaisierende Darstellung des Chemnitzer Lautsystems (KHAN / WEISE 2013). Auch das Wörterbuch der obersächsischen Mundarten (1984–2003) archiviert in vielen Teilen schwindenden oder gar geschwundenen Wortschatz. Die genannten dialektologischen Arbeiten beschäftigen sich deshalb nicht nur mit den Mundarten, sondern auch in wesentlichen Teilen mit der „obersächsischen Umgangssprache“, welche den Dialekt als

standardfernste Schicht ablöst, bzw. abgelöst hat. Die einzige größere jüngere empirische Arbeit zur Sprache in Sachsen, der Wortatlas der städtischen Umgangssprache (PROTZE 1997) berücksichtigt für die Wortschatzerhebung deshalb auch nicht mehr den Dialekt, sondern eben die Umgangssprache.

Die neuere Forschung legt dem entsprechend auch den Schwerpunkt nicht mehr auf die Basisdialekte sondern auf den Substandard bzw. das regionalsprachliche Spektrum, womit die Basisdialekte in größere Zusammenhänge eingebettet werden (HÜNECKE / KOCH / KOCH 2012, KEHREIN 2012, SCHAUFUB i. Dr., SCHAUFUB / SIEBENHAAR 2012). Im Rahmen der Erforschung der Sprachdynamik, der Erforschung der vertikalen und horizontalen Sprachschichtung (SCHMIDT / HERRGEN 2011), wird deutlich, dass sich in Dresden die Dialektalität in verschiedenen Kommunikationsformen nur wenig unterscheidet. KEHREIN bestätigt damit die früheren Befunde (FLEISCHER 1961), dass Dresdner Sprecher nicht mehr über basisdialektale Kompetenz verfügen, also keinen Dialekt als konsistentes Sprachsystem mehr realisieren können. Gleichzeitig realisieren die sie auch beim Vorlesen relativ viele standarddivergente Formen. Mit den mittelbairischen Sprechern aus Trostberg und den ostfränkischen aus Bamberg realisieren die Dresdener Sprecher beim Vorlesen mehr standarddivergente Formen als die Sprecher anderer Regionen. Die beiden Pole sprachlicher Verhaltensweisen stellen für sie folglich nur noch stilistische Varianten innerhalb einer Varietät dar. KEHREIN unterscheidet sie deshalb nur als Sprechlagen im Regiolekt und nicht als Varietäten. Diese geringe Variation ist für eine weitere Beschreibung der Sprache in Sachsen, und damit auch für Leipzig, prägend. Allerdings zeigt sich bei einer etwas breiter abgestützten Untersuchung im zentralen sächsischen Raum bei der selben Tendenz ein regional zumindest teilweise zu differenzierendes Bild (SCHAUFUB i. Dr.).

Da basisdialektale Laut- und Formenunterschiede in zentralen Teilen Sachsens kaum mehr strukturierend wirken, werden für eine Differenzierung verschiedener sprachlicher Regionen vermehrt suprasegmentale Aspekte genannt. Diese prägen auch die auch die standardnäheren Sprechlagen. Prosodische Unterschiede sind bislang in der Forschung erst großregional abgegrenzt worden (z. B. GILLES 2005, KÜGLER 2007, SELTING 2001, 2003). Diese prosodischen Besonderheiten spielen in der Wahrnehmung der Sprecher aus Sachsen selbst eine prominente Rolle und werden auch von Laien für eine kleinregionale Dialektgliederung als salient eingeschätzt (ANDERS 2010). Dieser Zusammenhang konnte bislang aber empirisch noch nicht gestützt werden. Dabei ist jedoch anzumerken, dass trotz der entsprechenden Beteuerungen der Sprecher eine regionale Zuordnung von Sprachaufnahmen innerhalb Sachsens kaum mehr gelingt (ANDERS 2010, SCHAUFUß i. Dr.).

Schwerpunkt dialektologischer Arbeiten ist – wie allgemein üblich – die lautliche Seite, die mit morphologischen und syntaktischen Aspekten ergänzt wird. Zusätzlich ist der dialektale Wortschatz im Wörterbuch der obersächsischen Mundarten (1984–2003) erfasst. Die Analyse der Laute erfolgt dabei weitgehend auf der Grundlage ohrenphonetischer Transkription bzw. der Laienschreibweise im Wenker Atlas. Seit den 1970er Jahren werden Dialektdaten auch instrumentalphonetisch analysiert. Die ersten Formantmessungen wurden von HEIKE / SCHINDLER (1970) mit Daten, die für Schlesischen Sprachatlas aufgenommen worden sind, vorgenommen, um die Transkripte aus dem Sprachatlas zu überprüfen. GÖSCHEL (1974) hat mit nordthüringischen Dialekten und solchen der erzgebirgischen Sprachinsel im Oberharz ebenfalls mitteldeutsche Dialekte untersucht, wobei die instrumentalphonetischen Analysen nicht die gewünschte Klärung phonologischer Fragen gebracht hatten. Seither sind einige wenige

instrumentalphonetische Arbeiten zur Bestimmung der Lautqualität deutscher Mundarten erschienen, die aktuellste (GACKSTATTER / NIEBUHR 2012) beschäftigt sich mit der Qualität der Langvokale in norddeutschen Mundarten. Entsprechende Verfahren sind für den Atlas of North American English (LABOV et al. 2006) in einem groß angelegten Unternehmen für 297 Städte angewendet und die Resultate sprachgeographisch ausgewertet worden. Weiterführende methodische Hinweise zur Vergleichbarkeit von Formantmessungen finden sich in aktuellen Einführungen in die Soziophonetik (THOMAS 2011, DI PAOLO / YAEGGER-DROR 2011). Für den dialektologischen Vergleich im Deutschen ist besonders VAN DER HARST (2011) bedeutsam, der verschiedene Mess- und Normalisierungsmethoden für die Unterscheidung der Vokalqualität und die Analyse des Vokalraums holländischer Mundarten vergleicht. Für die obersächsischen Mundarten liegen bislang keine entsprechenden instrumentalphonetischen Daten vor. Jedoch ist die ostmitteldeutsche Form der Standardaussprache von IIVONEN (1987 et passim) in mehreren Aufsätzen mit anderen regionalen Standardrealisierungen verglichen worden. Genuin instrumentalphonetisch sind die Arbeiten zur regionalen Prosodie, die jedoch alle nicht die (nicht mehr vorhandene) Mundart, sondern den Substandard behandeln. So hat GERICKE schon 1963 in einer der frühen Arbeiten zur regionalen Prosodie eine für Leipzig als typisch erachtete Kontur, den Fallbogen beschrieben. Diese ist später bei SELTING (2001) und GILLES (2005) auch für Dresden belegt worden und erweist sich im nationalen Vergleich als typisch ostmitteldeutsche Intonationskontur. Der gemäß GERICKE (1963) für Leipzig obligatorische Schluss auf halber Höhe ist aber für Dresden nicht bestätigt. KÜGLER (2007) zeigt in der Analyse der Akzentpositionierung, dass der Leipziger Tonakzent im Sinne der autosegmentalen Phonologie (GRICE / BAUMANN 2002) phonologisch dem Standard entspricht, allerdings

phonetisch steiler realisiert wird. SPIEKERMANN (2000) berücksichtigt für die Untersuchung des Silbenschnitts auch historische Daten aus Wurzen, einer Kleinstadt 30 km östlich von Leipzig. Dabei unterscheiden sich lange und kurze Vokale (sanft bzw. scharf geschnittene Vokale) weniger deutlich als in anderen Gebieten durch einen Dauerunterschied. Sie weisen dagegen einen stärkeren Qualitätsunterschied auf.

Die Stadt Leipzig selbst ist – trotz der Leipziger Schule – bislang wenig Schwerpunkt linguistischer Arbeit gewesen. 1881 hat ALBRECHT eine umfangreiche Ortsgrammatik und ein Wörterbuch der alten Stadtmundart herausgegeben. Dabei weist er nicht nur auf Unterschiede zwischen Stadt- und Bauernsprache, zwischen gezielter Sprechweise und der Sprache der Ungebildeten hin. Er bemerkt auch mehrfach den Rückgang der Mundart zugunsten von Formen, die der Schriftsprache näher stehen, was spätere Arbeiten für den gesamten zentralsächsischen Raum bestätigen. BAUMGÄRTNER (1959) hat in seiner Dissertation zur Syntax der Umgangssprache in Leipzig nicht mehr den Begriff Mundart verwendet. In seiner Arbeit untersucht er auch weniger die regionalen Varianten syntaktischer Variation als vielmehr generelle Aspekte von Mündlichkeit, die von der Norm der kodifizierten, an der Schrift orientierten Standardsprache abweichen. Auch die schon erwähnte Arbeit von GERICKE (1963), die als eine der ersten instrumentalphonetisch an regionalsprachliche Formen der Intonation herangegangen ist, geht nicht mehr vom Dialekt, sondern von der Umgangssprache aus.

Aktuelle Untersuchungen zur Leipziger Stadtsprache sind erst im Entstehen. Einen ersten Blick in die auch hier untersuchten Daten geben SCHAUFUß / SIEBENHAAR (2012). Dabei wird deutlich, dass die Sprecher selbst insgesamt vier Sprechlagen unterschieden, von denen sie zwei als Dialekt

bezeichnen und zwei als Hochdeutsch. Vom tiefen Dialekt distanzieren sich die meisten Sprecher, für einen Teil von ihnen ist der ‚gemäßigte Dialekt‘ eine Sprachform, die sie im Alltag verwenden und die sie positiv bewerten, in formellen Situationen wechseln sie zu standardnäheren Formen, beurteilen ihre eigene Standardkompetenz aber als eingeschränkt. Ein anderer Teil spricht sich selbst noch eine hohe Dialektkompetenz zu, versucht aber Dialektales zu vermeiden und spricht meist ein regional geprägtes Hochdeutsch. Eine letzte Gruppe sagt von sich, dass sie nur zwischen normgerechteren und normferneren Standardformen wechseln, Dialekt sprechen sie nicht und lehnen ihn ab. Diese Struktur widerspiegelt sich aber nur teilweise in den phonetischen Untersuchungen, denn auch die Sprecher, die sich noch eine hohe Dialektkompetenz zusprechen, können dialektale Formen (z. B. die Monophthongierung alter Diphthonge oder die Entrundung) nicht mehr systematisch verwenden, womit auch für Leipzig der Verlust dialektaler Kompetenz belegt ist. Im linguistischen Sinne ist das, was die Sprecher selbst noch als Dialekt bezeichnen, wegen der fehlenden Systematik also kein Dialekt mehr, sondern ein Substandard / Regiolekt mit noch unklaren Varietätenstatus. Dialektale Varianten einzelner Laute (dasselbe trifft auch für Morpheme und morphosyntaktische Besonderheiten zu, die noch stärker als die lautlichen abgebaut sind) kommen meist neben standardnäheren Varianten in unmittelbarer Umgebung vor. Die Wahl der Varianten einer linguistischen Variable ist demnach nicht mehr vom jeweils gewählten System abhängig sondern von einer mit außersprachlichen Faktoren / Funktionen korrelierenden Stilebene, wie das auch KEHREIN (2012) für Dresden feststellt. Diese Korrelation ist dabei jedoch nicht ursächlich zu verstehen in dem Sinne, dass eine bestimmte Situation einen Stil hervorbringt, wie das dem Labovschen Verständnis entspricht. Sie entspricht auch nicht einem erweiterten

Verständnis, dem Bells *audience design* zugrunde liegt, in dem der Sprecher die sprachliche Form realisieren will, von der er glaubt, dass der Hörer sie erwarte. Vielmehr bedingen sich sprachliche Stilebene und Situation gegenseitig, so dass nicht nur die Situation den sprachlichen Stil bestimmen kann, sondern auch der sprachliche Stil die Situation definieren kann. Ja, häufig werden dialektalere Formen im Gespräch bewusst zur Stilisierung eingesetzt, um Personen oder Gruppen zu kennzeichnen und um sich von diesen zu distanzieren, wie das BIRKNER / GILLES (2008) für die Dialektverwendung im Reality-Fernsehen oder ANDROUTSOPOULOS (2012) in youtube-Filmen und den dazugehörigen Kommentaren aufgezeigt haben. Aspekte der Stilisierung zeigen sich auch, wenn Gewährsleute in Sachsen versuchen Dialekt zu sprechen. Dabei werden oft nur einzelne Dialektmerkmale hervorgehoben. Es kann sogar vorkommen, dass dialektal intendierte Sprechteile nur prosodisch mit erhöhter Lautstärke und höherer Grundfrequenz (SCHAUFUß / SIEBENHAAR 2012) herausgestellt werden, ohne dass auf segmentaler Ebene Dialektmerkmale zu finden wären.

2 Entwicklung der Fragestellung

Die bisherige Forschung zur Sprache in Leipzig gibt also neben dem Befund aus dem Wenker-Atlas mit ALBRECHT (1881) fast gleichzeitig einen Blick in die Grammatik und den Wortschatz der mundartlichen Stadtsprache, der in vielen Bereichen sprachliche Variation dokumentiert. In den 1960er Jahren sind zwei Arbeiten zur Syntax bzw. zur Intonation der Leipziger Umgangssprache erschienen. Etwas neuer ist PROTZES Wortatlas der Umgangssprache (1997), der Leipzig miteinbezieht. Erste neuere

variationslinguistische Arbeiten bestätigen den in der Mitte des 20. Jh. dokumentierten Dialektabbau und weisen auf eine substandardsprachliche Strukturierung in der Wahrnehmung der Sprecher hin. An dieser Stelle setzt die vorliegende Studie ein. Während SCHAUFUß / SIEBENHAAR (2012) den Zusammenhang zwischen Spracheinschätzung und der Verwendung einzelner dialektaler Varianten herstellen, wird hier die lautliche Variation genauer betrachtet. Dabei geht es nicht um die Variation zwischen dialektaler und standardsprachlicher Varianten, wie z. B. der Monophthongierung alter Diphthonge oder der Lenisierung der Fortisplosive, die für die alte Mundart Leipzigs als charakteristisch gelten, sondern um die allgemeine Artikulationsbasis, d. h. die lautliche Qualität der einzelnen Vokale und deren Variation zwischen unterschiedlichen Sprechlagen. Da die Artikulation selbst außerhalb eines phonetischen Labors kaum zu messen ist, wird auf die messbare akustische Ebene ausgewichen, die mit der Artikulation zu einem hohen Grad korreliert, wie das in jeder Einführung in die Phonetik dargestellt wird. Es wird also die Frage gestellt, ob sich ein mundartlich intendierter Laut vom entsprechenden standardsprachlich intendierten Laut akustisch unterscheidet. Damit wird nicht eine allgemeine Dialektalität bzw. ein Abstand vom kodifizierten Standard auf der Basis einer ohrenphonetischen Transkription gemessen, wie das von HERRGEN / SCHMIDT (1985), KEHREIN (2012), LAMELI (2004) und anderen im Rahmen des REDE-Projekts vorgenommen wird, sondern einzelne Laute wie auch der gesamte Artikulationsraum werden auf ihre akustischen Parameter hin untersucht. Die Vokalrealisierung in unterschiedlichen Sprechlagen werden verglichen, welche durch die sprachliche Funktion voneinander abgegrenzt werden². Ein Vergleich

² Da der Status einzelner Abstufungen zwischen Standard-, Substandard- und Nonstandard-sprachen / Dialekten unklar ist, wird hier nicht von grammatisch unterschiedlichen Varietäten

mit dem kodifizierten Standard, bzw. der von Modellsprechern realisierten akustischen Form wird nicht vorgenommen, da die in der Literatur genannten Werte meist ebenfalls eine regionale Prägung ausweisen (IIVONEN 1987, 1994, MOOSMÜLLER 2007, SIMPSON 1998). Für die Berliner Analysen von SENDLMEIER / SEEBODE (o. J. [2006]) geben die Autoren an, dass keine dialektale Färbung vorhanden sei, was aber nicht durch Perzeptionstest überprüft wurde. Gleichzeitig sind Vergleiche durch unterschiedliche Erhebungsmethoden erschwert: IIVONEN (1987, 1994) und SENDLMEIER / SEEBODE (o. J. [2006]) untersuchen Wortlisten, die sich durch eine sehr hohe Artikulationsgenauigkeit auszeichnen. SIMPSON 1998 vergleicht wie die vorliegende Untersuchung Lesesprache und Spontansprache. Die umfassendste Untersuchung ist MOOSMÜLLER (2007), die Lese- und Spontansprache wie auch Wortlisten vergleicht.

Für den Untersuchungsraum ist lautliche Variation sowohl im Konsonantismus wie auch im Vokalismus zu finden. Die Veränderungen im Konsonantismus sind je ganz unterschiedlicher Art: Assimilationen, Apokopen und Synkopen in unterschiedlicher Kombination kommen vor: *sind* > *sin*, *ist* > *is*, *mal* > *ma*, *Kinder* > *Kinner*, *hinter* > *hinger*, *haben* > *ham*. Salient sind die Lenisierungen stimmloser Plosive *Torte* > *Dorde*, *packen* / *backen* > *baggn*, wobei die Lenisierung von *k* im Anlaut nach der älteren Literatur nur im Osterländischen verbreitet war: *Kabel* > *Gabel*. Auffällig sind ebenfalls die als pharyngal beschriebene Qualität von konsonantischem /ɾ/ (RUES et al. 2007: 93–97). Zudem findet sich die im Mitteldeutschen häufige Koronalisierung von /ç/ zu [ʃ] oder zu [ç] und eine im ganzen südlichen Teil des deutschen

mit pragmatischer, emischer und etischer Differenzierung ausgegangen, sondern von Sprechlagen als Verdichtungsbereichen variativer Sprachverwendung. Zur Unterscheidung von Varietäten und Sprechlagen vgl. LENZ 2003: 388–395.

Sprachgebiets verbreitete stimmlose Realisierung von /z/. Eine erweiterte Auflistung allgemein sächsischer standarddivergenter Merkmale findet sich in HÜNECKE et al. (2012: 150–151). Da die genannten konsonantischen Phänomene nur schwer in einen einzigen Zusammenhang zu bringen sind, konzentriert sich die vorliegende Arbeit auf den Vokalismus. Dieser kann im großen Ganzen mit drei artikulatorischen Parametern – Zungenhöhe, Zungenlage, Lippenspreizung/-rundung – und deren Kombination beschrieben werden. Die dialektologische Literatur hebt im Vokalismus neben der Monophthongierung alter Diphthonge besonders die Entrundung der gerundeten Vorderzungenvokale hervor. Die Entrundung ist zudem nicht ein isoliertes Phänomen, sie hat eine Zentralisierung des gesamten Vokalsystem zur Folge, wie das SCHMIDT (2012: 69–72) am Beispiel des ebenfalls umlautlosen Dialekts von Ludwigshafen hervorhebt. Gleichzeitig werden auch die hinteren Vokale etwas gespreizt und zentralisiert (siehe dazu auch die Darstellung zur Umgangssprache der Stadt Halle von WALLRAFF 2007: 35–36; 95–96; KHAN / WEISE 2013: 236 setzen für die Beschreibung des Chemnitzer Lautsystems sogar nur zentralisierte hintere Vokale an). Diese Zentralisierung der Hinterzungenvokale wird ebenfalls als typisch obersächsisches Merkmal wahrgenommen (PURSCHKE 2012: 275–293). Diese Zentralisierungstendenz, die in weiten Teilen des Obersächsischen vorkommt, wird auch von Laien erkannt und als undeutliche Aussprache, als Nuscheln u. ä. beschrieben (ANDERS 2010). Die Zentralisierung wird als kommunikativ salient eingeschätzt und trägt – vermutlich – zu einem größeren Teil zur negativen Einschätzung des Obersächsischen bei. Perzeptionsexperimente haben nachgewiesen, dass gerade diese Zentralisierungstendenz, die teilweise auch noch in der Vorleseaussprache vorkommt, wesentlich zur Einschätzung ostmitteldeutscher Tonaufnahmen als unkorrekt beiträgt (KEHREIN 2012: 331).

Für den alten Dialekt geben BECKER / BERGMANN (1969) auf lautlicher Ebene zudem großräumig Hebungen von mhd. ê, ô, o, Senkung von mhd. ë und viele kleinräumige Besonderheiten an, die aber alle schon als im Rückgang befindlich oder gar als geschwunden beschrieben werden.

Ob in Leipzig wie in Dresden ein geringer Unterschied zwischen verschiedenen Erhebungssituationen vorliegt, wie KEHREIN (2012) festgestellt hat, soll an dieser Stelle überprüft werden. Dazu werden unterschiedliche stilistischen Ausprägungen anhand verschiedener Erhebungssituationen untersucht. Die Forschungsfrage wird nun eingeschränkt: Unterscheiden Leipziger ihre Vokale in verschiedenen Sprechsituationen? Und wenn sie das tun, wird gefragt, wie sich der Artikulationsraum und die einzelnen Vokale unterscheiden, die in verschiedenen Sprechsituationen von Leipzigern realisiert werden?

3 Methodisches

3.1 Datenerhebung und -auswahl

Um Daten zu unterschiedlichen Sprechstilen zu erhalten, wurden im Rahmen eines Seminars an der Universität Leipzig von Studierenden insgesamt 22 Gewährspersonen in verschiedenen Situationen aufgenommen. Die in der Stadt Leipzig lebenden und da aufgewachsenen Gewährspersonen im Alter zwischen 30 und 87 Jahren, die im Wesentlichen einer bildungsnahen Mittelschicht zuzuordnen sind, gaben Auskunft über ihre Einschätzung verschiedener Sprachlagen und deren Verwendung in der Stadt Leipzig. Zudem haben sie alle vor dem Interviews einen Teil der Wenkersätze, die für das Ostmitteldeutsche

relevante Merkmale aufweisen, standardsprachlich gelesen und nach dem Interview in ihre tiefste mundartliche Sprechlage übersetzt. Diese letzte Aufgabe war für einige der Sprecher mit erheblichen Schwierigkeiten verbunden, da sie sich entweder selbst keine Mundartkompetenz mehr zusprachen oder die Verwendung standardferner Formen außerhalb des vertrauten Kreises als unangebracht betrachteten. Diese Haltung sowie die schriftliche Vorlage und die Anwesenheit einer fremden Person haben es vielen Gewährsleuten erschwert ihre standardfernste Sprachlage zu aktivieren. Die Übersetzung der Wenkersätze kann damit nicht als authentisch dialektal bewertet werden, sondern als intendierter Dialekt. Dieser intendierte Dialekt gibt damit nicht eine reale Sprachverwendungssituation wieder, sondern ist als Ausdruck mundartlicher Kompetenz zu betrachten. Mit dieser Einschätzung liegen für jede Gewährsperson drei Aufnahmen unterschiedlicher Sprechlagen vor, die sich im Allgemeinen, wie in SCHAUFUß / SIEBENHAAR (2012) beschrieben, durch unterschiedliche Anteile obersächsischer Marker unterscheiden. Für einen Teil der Gewährsleute liegen auch noch Freundesgespräche (vgl. LENZ 2003, KEHREIN 2012) vor, die aber noch nicht ausgewertet sind. Die Aufnahmen wurden mit einem Olympus LS 11-Gerät in der Wohnung der Gewährspersonen aufgenommen. Sie erreichen damit bei Weitem nicht Studioqualität, zudem konnten Nebengeräusche mit diesem Aufnahmesetting nicht vermieden werden, sie beeinträchtigten die phonetische Analyse im Allgemeinen aber nicht, so dass die Aufnahmen im großen Ganzen für instrumentalphonetische Analysen genutzt werden. In Einzelfällen mussten einige kürzere Passagen von der Analyse ausgeschlossen werden.

Für die vorliegende Analyse wurden schließlich nur Aufnahmen von Frauen ausgewählt. Diese haben ein Alter zwischen 30 und 75 Jahren bei einem arithmetischen Mittel von 54.3 Jahren. Die Einschränkung auf Frauen

hat mehrere soziolinguistische und akustische Gründe. In soziolinguistischer Hinsicht wird seit Jahrzehnten auf Unterschiede zwischen Männern und Frauen hingewiesen. In der amerikanischen Linguistik wurde immer wieder betont, dass Frauen aus der Mittelschicht besonders in formellen Situationen stärker zu prestigereicheren Varianten tendieren als Männer. Allerdings sind die Unterschiede im deutschen Sprachraum nur teilweise bestätigt worden. Mehrere aktuelle variationslinguistische Arbeiten haben sich deshalb – und auch aus datenökonomischen Gründen – darauf beschränkt, nur ein Geschlecht zu untersuchen, wobei meist nur Männer untersucht wurden (LENZ 2003, LAMELI 2004, KEHREIN 2012, anders GACKSTATTER / NIEBUHR 2013). Um die soziolinguistische Varianz zu reduzieren, wird auch in der vorliegenden Untersuchung nur ein Geschlecht analysiert, allerdings sollen diesmal die Frauen sprechen. Dieser Entscheid ist aber auch in akustischer Hinsicht sinnvoll: Frauen und Männer haben physiologisch einen unterschiedlich großen Artikulationsraum, wodurch sich nicht nur die Grundfrequenz, sondern auch die Werte der Formanten zwischen den Geschlechtern unterscheiden. Dieser Unterschied ist immer wieder festgestellt worden, allerdings ist die Frage, ob dies nur durch physiologische oder auch durch soziale Faktoren begründet sei, bislang nicht abschließend beantwortet (siehe dazu VAN DER HARST 2012: 47–48). Mit verschiedenen Normalisierungsverfahren können die auf dem Faktor Geschlecht beruhenden Messwertdifferenzen zumindest teilweise ausgeglichen werden. Die Einschränkung auf Frauen erlaubt es aber auch, nicht normalisierte Werte zu vergleichen. Bei der Normalisierung einer einheitlicheren Gruppe erscheinen zudem weniger Artefakte, so dass die individuellen Unterschiede deutlicher zutage treten.

3.2 Datenaufbereitung

3.2.1 Auswahl und Messung der Formantwerte

Die Aufnahmen wurden transkribiert und im Phonetikprogramm Praat (BOERSMA / WEENINK 2013) auf der Basis des Oszillogramms und des Sonogramms manuell gelabelt. Dabei wurden diejenigen Vollvokale markiert, die nicht durch Nebengeräusche, Knarrstimme, Oktavsprünge oder offensichtliche Messfehler, z. B. durch fehlende oder zusammenfallende Formantwerte, gestört waren. Ausgeschlossen wurden auch Vokale, welche durch Koartikulationserscheinungen nicht deutlich zu segmentieren waren. Dieses letzte Kriterium ist selbstredend sehr problematisch, da dies keine objektiven Faktoren zum Ausschluss einzelner Laute zur Verfügung stellt. Es führt auch dazu, dass kurz realisierte Vokale eher von der Analyse ausgeschlossen werden. Da gerade die kürzeren Vokale stärker zur Zentralisierung tendieren, weil weniger Zeit für das Erreichen der ‚korrekten‘ Zungenposition zur Verfügung steht, tragen die Analysen einen leichten Bias dahingehend, dass eher mehr peripherere Vokale gemessen werden als tatsächlich realisiert werden (siehe dazu aber auch Abschnitt 5.3).

Die Messung der Formanten erfolgt automatisiert mit Hilfe eines Skripts in Praat. Für die Formantanalyse wurde die LPC-Analyse mit der Erfassung von fünf Formanten im Spektrum vom 0–5500 Hz, wie das für die Analyse von Frauenstimmen empfohlen wird (BOERSMA / WEENINK 2013, Manual), im Abstand von 10 ms mit einer Fenstergröße von 25 ms gewählt. Für die Auswertung der Monophthonge wurden zwei Messpunkte um die Mitte des Vokals, die 45 %- und 55 %-Position innerhalb des Vokals, berücksichtigt. Für Monophthonge wird in der Mitte die stabilste und deutlichste Lautqualität

angenommen (HARRINGTON 2010a: 85). Da, wo sich diese beiden Messungen um mehr als 10 % unterscheiden, wird davon ausgegangen, dass ein Messfehler vorliegt. Die entsprechenden Laute wurden von der Analyse ausgeschlossen, genau so wie diejenigen, bei denen nur ein Messwert an beiden Messpunkten vorliegt. Zudem wurden die Häufigkeitsverteilung des Frequenzspektrums für die beiden ersten Formanten jedes Vokals überprüft und Outliner nach einer manuellen Überprüfung aus der weiteren Analyse ausgeschlossen³. Messfehler der automatischen LPC-Analyse konnten so nicht ganz sicher erkannt, aber weitgehend minimiert werden. Mit einem solchen Vorgehen ist aber auch nicht zu vermeiden, dass korrekte Messungen fälschlicherweise aus der Analyse ausgeschlossen werden. Beide Fehlertypen können mit teilautomatisierten Verfahren nicht mit absoluter Sicherheit ausgeschlossen werden, die manuelle Korrektur führt aber andere Fehlerquellen ein, so dass darauf verzichtet wurde.

Das artikulatorisch-akustische Modell, das diesem Verfahren zugrunde liegt, suggeriert, dass Monophthonge eine konstante Lautqualität aufweisen, was jedoch nur in den wenigsten Fällen zutreffend ist (vgl. VAN DER HARST 2012; GACKSTATTER / NIEBUHR 2013). Es ist jedoch eine übliche Vereinfachung, die für eine artikulatorische Extremposition der Monophthonge und damit den Vergleich der Vokalpositionen und die Berechnung des Artikulationsraumes meist ausreichend ist (HARRINGTON 2010b: 122, THOMAS 2011: 152).

³ Anzumerken ist hier, dass sich sämtliche statistischen Outliner als fehlerhafte Messungen des formant-tracking-Algorithmus erwiesen haben. Manuelle Korrekturen der automatisch berechneten Formantwerte, wie beispielsweise von MOOSMÜLLER (2007: 25–28) oder SIMPSON (1998: 48–49) eingesetzt, wurden nicht vorgenommen. Stattdessen wurden die entsprechenden Vokale von der weiteren Analyse ausgeschlossen.

3.2.2 Grundlagen der Analyse, Normalisierung

Die Formantwerte werden für die folgende Analyse teilweise als Rohwerte, teilweise als normalisierte Werte berücksichtigt. Beide Verfahren zeigen Vor- und Nachteile, die für die jeweiligen Analysen gegeneinander abgewogen werden. Bei nicht normalisierten Daten werden die Unterschiede, gemessen in Hz, zwischen den Sprecherinnen beibehalten und können damit untersucht werden. Das ist aber insofern problematisch als anatomische Unterschiede zwischen den Sprecherinnen sich auch als unterschiedliche Vokalqualitäten und unterschiedlich große Artikulationsräume (siehe 3.3) manifestieren, was aber nur als akustischer, nicht als linguistischer Unterschied zu werten ist. Unterschiede in den höheren Frequenzen erscheinen zudem größer als in den tieferen Frequenzen. Dieser Überinterpretation akustischer Merkmale kann mit einer Normalisierung teilweise begegnet werden, da sie interindividuelle Unterschiede ausgleicht. Allerdings birgt auch jede Normalisierung Nachteile oder geht von Voraussetzungen aus, die mit der vorliegenden Untersuchung nicht erfüllt werden können: Das von LABOV et al. (2006: 40) verwendete Verfahren ist beispielsweise erst bei großen Datenmengen – mehr als 345 Sprecher – stabil und ist dem entsprechend für die vorliegende Untersuchung ungeeignet. Die Lobanov-Normalisierung (statistisch eine z-Transformation) hat sich für Sprachvergleiche als häufig geeignet erwiesen (THOMAS 2011, VAN DER HARST 2011), sie wird deshalb auch hier für einzelne Darstellungen verwenden. Allerdings hat sie den Nachteil, dass sie durch unterschiedliche ausgeprägte Vokalsysteme beeinflusst wird (THOMAS / KENDALL 2007–2010, THOMAS 2011: 168), was gerade das untersuchte Ziel beeinträchtigt. Zudem wird sie durch unterschiedlich große Datensätze, wie sie hier durch die

verschieden langen Teile der Aufnahmen und die unterschiedlich langen Interviews gegeben sind (vgl. Tabelle 1 in Abschnitt 4), beeinflusst. VAN DER HARST (2011) schränkt deshalb für seinen Vergleich holländischer Dialekte die untersuchte Datenmenge ein, was die Reliabilität beeinträchtigt; eine solche Einschränkung kann bei eher kleineren Datensätzen, wie dem vorliegenden, aber ebenfalls zu Verzerrungen führen. Weil also sowohl zur Analyse mit nicht normalisierten Daten, wie auch zur Analyse normalisierter Daten Bedenken vorgebracht werden, werden die Daten mit beiden Verfahren analysiert. Für die Präsentation wird nur je auf ein Verfahren zurückgegriffen, die Wahl des jeweils verwendeten Verfahrens wird da je kurz begründet.

Die Normalisierung erfolgt mit einer z-Transformation, wie sie von LOBANOV (1971) vorgeschlagen worden ist. Das Verfahren hat sich in mehreren Vergleichsuntersuchungen als stabil erwiesen (THOMAS 2011: 165–170, VAN DER HARST 2011). Für die z-Transformation werden die vorhandenen Werte für jeden Formanten so transformiert, dass das arithmetische Mittel 0 entspricht und die Varianz 1⁴. Im Zuge dieses Lobanov-Verfahrens werden alle Vokale bzw. alle Messungen für jeden einzelnen Sprecher berücksichtigt, es wird damit als vokal-extrinsische Methode bezeichnet. Da die Daten für jeden einzelnen Sprecher normalisiert werden, gilt es als sprecher-intrinsisches Verfahren. Mittelwert und Varianz entsprechen also für jeden Sprecher 0 bzw. 1, obwohl sich die Rohwerte unterscheiden. Für die vorliegenden Aufnahmen hat der erste Formant beispielsweise bei einer Sprecherin den Mittelwert über alle Aufnahmen und Vokale von 515 Hz, während er bei einer anderen

⁴ THOMAS / KENDALL (2007–2010) machen darauf aufmerksam, dass für das Verfahren heute nicht mehr die rms, die quadratische Abweichung vom Mittelwert, sondern die Standardabweichung verwendet wird, was die Resultate aber nicht beeinflusst. Für die vorliegende Untersuchung wird entsprechend der aktuellen Forschung ebenfalls die Standardabweichung verwendet.

Sprecherin bei 596 Hz liegt. Für die normalisierten Daten liegt der Mittelwert definitionsgemäß für beide Sprecherinnen bei 0. Die Umrechnung der Formantwerte für jeden einzelnen Vokal auf z-normalisierte Werte erfolgt nach der folgenden Formel:

$$F_{[V]}^N = \frac{(F_{[V]} - \bar{F})}{s}$$

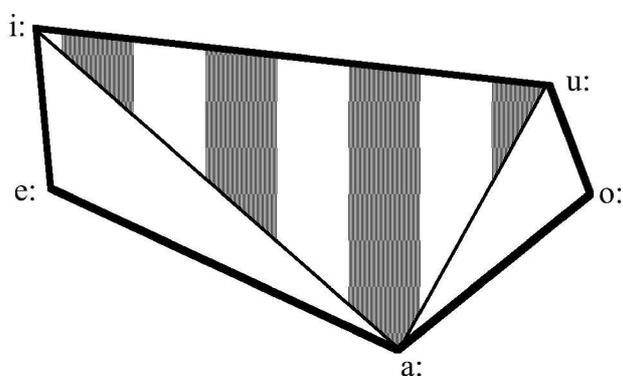
Dabei ist $F_{[V]}^N$ der normalisierte Wert für einen Formanten eines bestimmten Vokals. $F_{[V]}$ ist der gemessene Formantwert des Vokals, \bar{F} ist der Mittelwert des Formanten für alle Messungen und s ist die Standardabweichung.

3.2.3 Berechnung der Größe des Artikulationsraumes

Für den Vergleich der Artikulationsbasis in den verschiedenen Erhebungssituationen wird auch die Größe des jeweiligen genutzten Artikulationsraumes berechnet, und zwar für jede Erhebungssituation und für Lang- und Kurzvokale gesondert. Als Grundlage dienen die für jeden Vokal gemittelten Formantwerte bzw. die normalisierten Werte. In vielen Fällen wird die Artikulationsbasis mittels der Fläche des Dreiecks durch die Eckpunkte [i: – u: – a:] bestimmt, wie das in Abbildung 1 mit der grau hinterlegten Fläche dargestellt wird. JACEWICZ et al. (2007: 1456) und VAN DER HARST (2011: 241–243) kritisieren dieses Verfahren, weil so nur ein Teil des genutzten Vokalraums berücksichtigt wird. Bei der Messung der Lautqualität deutscher Vokale in intendierter Standardlautung (IIVONEN 1987, 1994, MOOSMÜLLER 2007, SIMPSON 1998: 60, SENDLMEIER / SEEBODE o. J. [2006], siehe dazu auch

Abb. 5) und auch in der vorliegenden Untersuchung ist bei den meisten Sprecherinnen nicht [u:] sondern [o:] der hinterste Laut, und [e:] wird häufig fast ebenso weit vorne realisiert wie [i:]. Deshalb entspricht eine solche Flächenberechnung nicht dem tatsächlich im Deutschen realisierten Artikulationsraum.

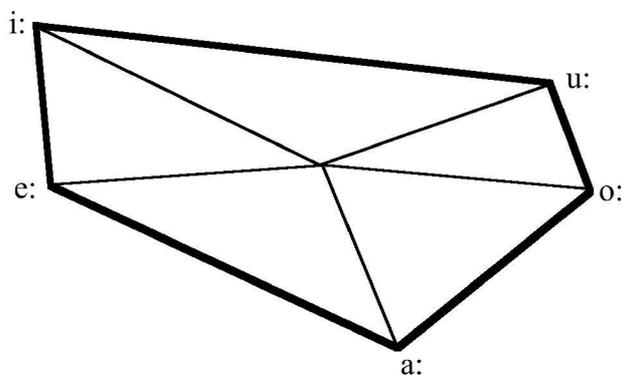
Abbildung 1: Schematische Darstellung der häufig verwendeten Berechnung des dreieckigen Vokalraumes (in grau)



Um eine adäquatere Messung zu erreichen, wurde eine Berechnung gewählt, die den ganzen Vokalraum ausmisst. Dafür wurden – wie in Abbildung 2 für die Langvokale dargestellt – jeweils die Flächen der Dreiecke zwischen benachbarten Vokalen und einem zentralen Punkt (500Hz / 1500Hz) für die Rohwerte bzw. zum 0 / 0-Wert bei den normalisierten Werten berechnet. Die Flächen der fünf Dreiecke wurden dann zusammengezählt. Der Langvokal /ɛ:/ ist nur selten realisiert worden, womit nicht für alle Sprecher in allen Erhebungssituationen ein Beleg vorliegt. Da, wo /ɛ:/ realisiert worden ist, liegt zudem eine sehr hohe Variation vor, weil /ɛ:/ die Tendenz hat, mit /e:/ zusammenzufallen (im Standard: KRECH u. a. 2009: 58; für die Leipziger Mundart: ALBRECHT 1881: 7; für den obersächsischen Raum HÜNECKE et al.

2012: 163). /ɛ:/ wurde so mit unterschiedlicher Qualität als [e:] oder [ɛ:] realisiert, womit auch dessen Phonemstatus in Frage gestellt ist. Die geringe Datenbasis würde also den Wert von [ɛ:] als sehr unsicher erscheinen lassen. Weil besonders nicht für alle Sprecherinnen in allen Erhebungssituationen /ɛ:/ vorliegt, würde das zufällige Vorhandensein bzw. Fehlen dieses Lautes die gemessene Größe des Vokalraumes stark beeinflussen, so dass ein Vergleich der Artikulationsräume nicht mehr sinnvoll möglich wäre. All das spricht dagegen, [ɛ:] in die Berechnung einfließen zu lassen, weshalb es für die gesamte Analyse außen vor gelassen wurde. Es ist aber nicht von der Hand zu weisen, dass damit bei einzelnen Sprecherinnen nicht der ganze Artikulationsraum mitberücksichtigt wird.

Abbildung 2: Schematische Darstellung der hier verwendeten Größenberechnung des Vokalraumes



Die Berechnung der Dreiecksfläche erfolgt nach der Formel von Heron:

$$\text{Dreiecksfläche} = \sqrt{s(s-a)(s-b)(s-c)}$$

$$\text{wobei } s = \frac{a+b+c}{2}$$

wobei a jeweils die Strecke zwischen zwei Vokalen darstellt und b und c jeweils die Strecke zwischen den Vokalen und dem Mittelpunkt bei 500Hz / 1500Hz bzw. dem 0 / 0-Punkt z-normalisierter Werte.

Die Gesamtartikulationsfläche ist die Summe der fünf Dreiecke.

4 Datengrundlage

Von den insgesamt 22 Leipziger Aufnahmen waren 16 von Frauen, daraus wurden die 12 mit der besten Aufnahmequalität für die Analyse ausgewählt. Die Interviews inklusive der Lektüre und Übersetzung der Wenkersätze dauern jeweils zwischen 17 und 38 Minuten.

Tabelle 1 gibt die Anzahl gemessener Laute und die Anzahl der für die Analyse verwendeten Laute wieder. Die gerundeten Vorderzungenvokale sind extra aufgeführt, da sie im Basisdialekt ja eigentlich nicht vorkommen (ALBRECHT 1881: 7–10). Allerdings findet sich schon im Wenker Atlas ein Nebeneinander gerundeter und ungerundeter Varianten, sie wurden auch hier von allen Sprecherinnen in allen Situationen realisiert, auch in der intendierten Ortsmundart, wobei diese da oft neben der entrundeten Variante stehen, wodurch die Datenbasis dünn wird. Für die Berechnung des Vokalraumes (siehe 3.2.3) wurden die gerundeten Vorderzungenvokale deshalb nicht berücksichtigt. Weil deren zweiter Formant im Vergleich zum zweiten Formanten der ungerundeten Vokal abgesenkt ist, liegen sie auch mit wenigen Ausnahmen, in denen /y:/ höher als /i:/ realisiert wird, innerhalb des Vokaltrapezes der primären Kardinalvokale. Ebenfalls nicht berücksichtigt sind die Diphthonge. Von der Analyse ausgeschlossen wurden diejenigen Vokale, deren Formantwerte nach Abschnitt 3.2.1 als unsicher eingestuft wurden.

Tabelle 1: Datenbasis nach Lautklassen

Laut	Kurzvokal (ohne gerundete vorne)	Kurzvokal (gerundet vorne)	Langvokal (ohne gerundete vorne)	Langvokal (gerundet vorne)	Diphthong
gesamt	4334	276	3330	332	1962
ausgeschlos- sen (unklare Messdaten)	531	21	586	29	
analysiert	3803	-	2744	-	-

Tabelle 2 gibt die Anzahl analysierter Monophthonge nach Gewährsperson und Vokaltyp wieder. Tabelle 3 zeigt die Aufteilung nach Person und Erhebungssituation. Dabei wird deutlich, dass trotz gleicher Textvorlage bei den Wenkersätzen nicht bei allen Sprecherinnen gleich viele Vokale analysiert werden können. Da teilweise ungenaue Realisierungen oder Messfehler vorliegen, mussten sie von der Analyse ausgeschlossen werden. Für alle Sprecherinnen liegen aber für jeden Laut und jede Erhebungssituation (außer [ɛ:], der nicht berücksichtigt wurde) mindestens zwei Belege vor.

Tabelle 2: Analysierte Monophthonge nach Lautklasse und Gewährsperson

Gewährsperson	Kurzvokal	Langvokal	Total
IS36	319	251	570
GR40	456	300	756
AS42	459	369	828
GN41	248	190	438
AK45	375	263	638
AL51	339	191	530
IL55	243	135	378
GL57	327	291	618
GA66	246	218	464
AR81	245	167	412
IT81	227	132	359
CL85	319	175	494
	3803	2682	6485

Tabelle 3: Analysierte Monophthonge nach Erhebungssituation und Gewährsperson

Gewährsperson	Wenkersätze intendierter Standard	Interview	Wenkersätze intendierter Dialekt	gesamt
IS36	95	387	88	570
GR40	89	568	99	756
AS42	92	664	72	828
GN41	84	267	87	438
AK45	101	422	115	638
AL51	73	413	44	530
IL55	72	228	78	378
GL57	82	484	52	618
GA66	69	300	95	464
AR81	60	276	76	412
IT81	65	232	62	359
CL85	62	363	69	494
Total	944	4604	937	6485

5 Analysen

5.1 Die Strukturierung des Vokalraums

5.1.1 Vergleich der Erhebungssituationen

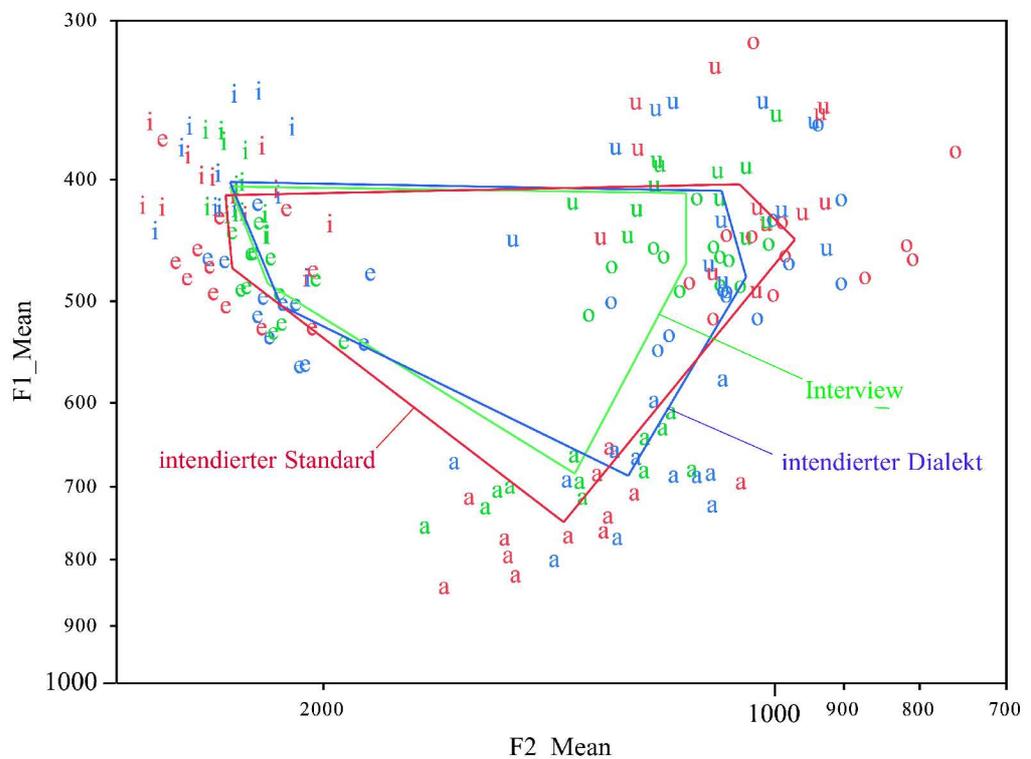
Abbildung 3 und 4 zeigen die nicht normalisierten Mittelwerte der Realisierungen für jeden Laut jeder Sprecherin in jeder Erhebungssituation. Zudem sind die Mittelwerte aller Laute für alle Sprecherinnen und jede Erhebungssituation mit Linien ausgezeichnet, um den je Erhebungssituation genutzten Vokalraum zu verdeutlichen. Der Vergleich der einzelnen Vokale zeigt deutlich die relativ große Variation.⁵

Bei den Langvokalen (Abbildung 3) ist einzig [a:] ohne jegliche Überschneidungen von den anderen Vokalen abgegrenzt. Die Vorderzungenvokale [i:] und [e:] zeigen einzelne Überschneidungen, besonders in den Fällen, in denen gleichzeitig eine Zentralisierung zur Mitte hin vorliegt. Die Hinterzungenvokale [u:] und [o:] sind dagegen nur schwer voneinander zu trennen; beide weisen sowohl für F_1 als auch für F_2 eine große Streuung auf. Hinterzungenvokale und Vorderzungenvokale sind trotz der für das Sächsische immer wieder genannten Zentralisierung aber in allen Fällen deutlich getrennt. [i:] und [u:] einerseits und [e:] und [o:] andererseits sind jeweils durch sehr

⁵ Die normalisierte und nicht normalisierten Versionen unterscheiden sich nur geringfügig, besonders die Hinterzungenvokale sind in der normalisierten Version trennschärfer. Da der Unterschied aber nicht besonders groß ist, wird an dieser Stelle nur die eine Version dargestellt und analysiert. Dadurch, dass in der nicht-normalisierten Version die Daten in Hertz wiedergegeben sind, lassen sich die Daten einfacher nachvollziehen und mit anderen Analysen vergleichen, weshalb dieser Version in diesem ersten Überblick Vorrang gegeben wurde.

deutliche Unterschiede auf dem zweiten Formanten voneinander geschieden. Die in der Analyse nicht berücksichtigten gerundeten Vorderzungenvokale [y:] und [ø:] nehmen mit F₂-Werten von 1498–2008 Hz diese zentrale Position ein.

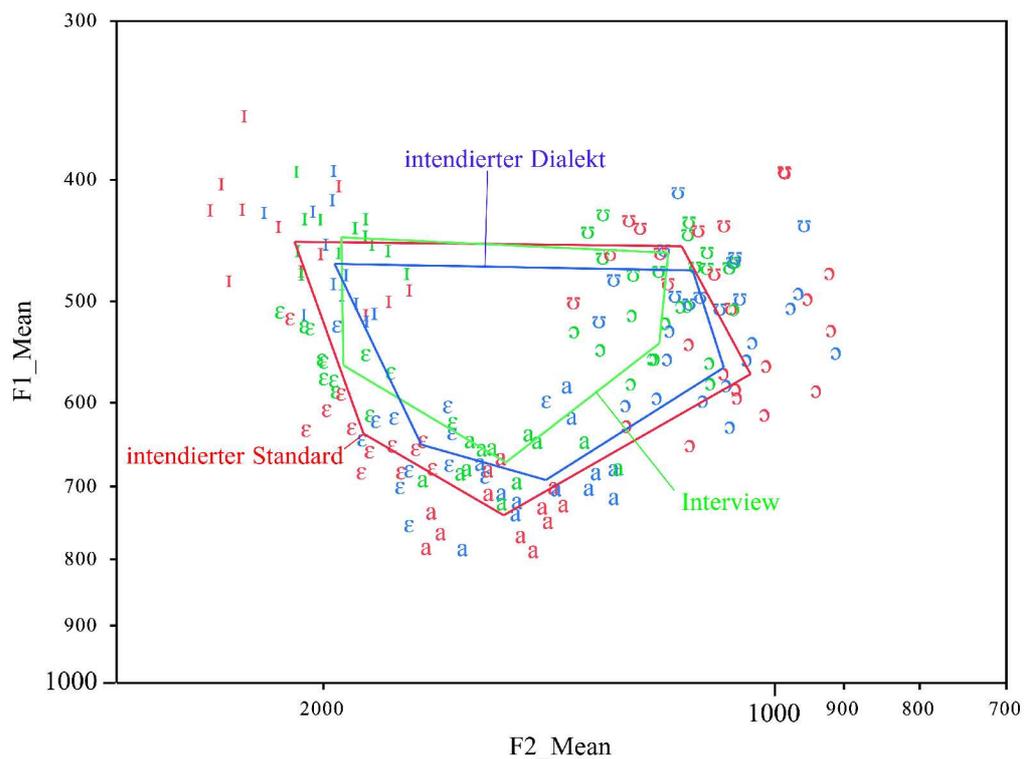
Abbildung 3: Vokalraum der Langvokale, Belege für die einzelnen Sprecherinnen und Auszeichnung des Vokalsraums für die einzelnen Erhebungssituationen



Diese Trennung zwischen Vorder- und Hinterzungenvokalen wird auch bei den Kurzvokalen (Abbildung 4) aufrecht erhalten, obwohl Kurzvokale auf den ersten Blick erkennbar mehr streuen als die Langvokale und obwohl der Artikulationsraum merklich verkleinert ist. Dadurch dass [ɛ] signifikant tiefer realisiert wird als [e:] ist die Abgrenzung zu [a], das im Vergleich zum

Langvokal auch etwas weiter vorne realisiert wird, weniger deutlich. Auffällig ist bei den kurzen Vorderzungenvokalen besonders die große Streuung auf dem ersten Formanten, der mit der Zungenhöhe korreliert. Bei den Hinterzungenvokalen unterscheidet sich [ɔ] deutlich von [a]. Die Laute [ɔ] und [ʊ] zeigen dagegen einen großen Überschneidungsbereich. Wie bei den Langvokalen ist die Streuung bei den Hinterzungenvokalen sowohl auf dem ersten wie auch auf dem zweiten Formanten groß.

Abbildung 4: Vokalraum der Kurzvokale, Belege für die einzelnen Sprecherinnen und Auszeichnung des Vokalsraums für die einzelnen Erhebungssituationen



Werden die gemittelten Vokalräume der einzelnen Erhebungssituationen verglichen – die ausgezogenen Linien in den Abbildungen 3 und 4 – so treten bei Kurz- und Langvokalen deutliche Unterschiede zutage. Bei den Langvokalen (Abbildung 3) fällt auf, dass sich die Höhe der Hochzungenvokale zwischen den verschiedenen Erhebungssituationen kaum unterscheidet, was ähnlich bei (MOOSHAMMER / GENG 2008) im Vergleich akzentuierter und nicht akzentuierter Vokale in experimentellem Setting belegt ist. Der Vokal [i:] wird zudem auch in der Zungenlage kaum variiert und stellt damit im gesamten Vokalraum eine Konstante dar. Gegenüber der standardintendierten Leseaussprache wird [e:] im intendierten Dialekt und im Interview gesenkt und zentralisiert. Die Realisierungen von intendiertem Dialekt und Interview unterscheiden sich kaum. Das Resultat ist durch die instrumentalphonetische Messung wohl etwas differenzierter als die ohrenphonetischen Arbeiten von WALLRAFF (2007: 89) für Halle und KEHREIN (2012: 229) für Dresden, die beide zwischen den Erhebungssituation keine Unterschiede feststellen. Die Hinterzungenvokale erscheinen im Interview und etwas weniger deutlich im intendierten Dialekt gegenüber der Leseaussprache mit einem tieferen F₂-Wert als zur Mitte hin zentralisiert – auch dies scheint für Hinterzungenvokale die Hauptrichtung der Zentralisierung im Deutschen zu sein (MOOSHAMMER / GENG 2008). Im Vergleich zu den Analysen des norddeutsch geprägten Kiel Korpus ist die Reduktion besonders der Hinterzungenvokale stärker ausgeprägt (SIMPSON 1998: 89–94). [o:] wird in beiden Korpora gegenüber der standardintendierten Leseaussprache gesenkt. Im Vergleich zu SIMPSON (1998: 89–94) fallen zwei weitere Aspekte auf, die als regionalsprachlich zu bewerten sind: Einerseits wird [a:] in der spontansprachlichen Realisierung nach hinten verschoben oder allenfalls gerundet, da ein tieferer F₂-Wert auch mit einer Rundung einhergeht, während

bei SIMPSON eine Vorverlagerung zu beobachten ist. Auf der anderen Seite ist die starke Hebung von [ɛ] bemerkenswert, die bei SIMPSONS Analyse des Kieler Korpus nicht zu beobachten ist.

Im Gegensatz zur Darstellung bei KEHREIN (2012: 228) für Dresden lässt sich somit bei den untersuchten Leipziger Sprecherinnen ein etwas deutlicherer Unterschied zwischen den einzelnen Erhebungssituationen feststellen. Die Zentralisierung der Hinterzungenvokale ist zudem im Interview ausgeprägter als im intendierten Dialekt. Damit wird die Zentralisierung der Hinterzungenvokale von den Sprecherinnen selbst weniger als dialektales Element gestaltet, sondern ist eher als Zeichen für Informalität zu werten. Das ist insofern erstaunlich, da diese Zentralisierung im überregionalen Vergleich und in Perzeptionsexperimenten als stigmatisiertes (HERRGEN / SCHMIDT 1985) und typisch regionalsprachliches Merkmal des Sächsischen wahrgenommen (KEHREIN 2012: 228, WIESINGER 1983b: 1104, KHAN / WEISE 2013: 236) wird. [a:] wird im intendierten Dialekt und im Interview gehoben, im intendierten Dialekt wird es auch deutlich nach hinten verschoben. Dass spontansprachliche Vokale im Vergleich zur Leseaussprache zentralisiert werden, hat sich in mehreren Untersuchungen bestätigt (GENDROT / ADDA-DECKER 2005, NICOLAIDIS 2003, SIMPSON 1998, WREDE u. a. 2000). Die damit verbundene Verkleinerung des Artikulationsraumes hat aber vermutlich nicht nur mit der Kürze der Laute zu tun⁶, sondern ist wohl ein stilistisches Merkmal der Spontansprache, welches in den vorliegenden Daten bei den Hinterzungenvokalen mit der kürzeren Vokaldauer korreliert. Diese

⁶ Die von LINDBLOM 1963 formulierte erste Fassung der undershoot-Hypothese, die voraussagt, dass kürzere Vokale zentralisierter seien als längere, konnte trotz ihrer Einfachheit und auf den ersten Blick einleuchtenden Ansatzes leider nicht durchgehend bestätigt werden (vgl. HARRINGTON 2010a: 91–95 und die da zitierte Literatur). Vielmehr sind Koartikulation, Häufigkeit und Sprechstil Faktoren, die einen wesentlichen Einfluss haben. Siehe dazu auch Abschnitt 5.2.

Korrelation zwischen Vokaldauer und F_2 ist dem entsprechend für [o:] und [u:] signifikant, wenn auch eher gering ($r = 0.238^*$ für [o:] bzw. $r = 0.272^*$ für [u:]). Dass [o:] und [u:] aber im intendierten Dialekt zentralisiert werden, passt nicht in dieses Bild. Denn während ein standardintendiertes [o:] 114 ms lang ist und ein [o:] im Interview nur 95 ms, was die Zentralisierung im Interview erklären kann, so ist das [o:] im intendierten Dialekt mit 140 ms sogar deutlich länger als in der Leseaussprache. Die Zentralisierung kann im intendierten Dialekt also nicht durch die Vokallänge, bzw. durch deren Kürzung gegenüber der Leseaussprache erklärt werden. Es muss deshalb davon ausgegangen werden, dass das [o:] der intendierten Mundart auch tatsächlich eine andere (zentralisiertere) Qualität haben soll, als das der intendierten Standardsprache. Dasselbe gilt für die Hebung von [a:], die im intendierten Dialekt zusammen mit einem Backing also als bewusst nach oben zentralisierte Realisierung erscheint. Die beiden Argumentationsstränge scheinen auf den ersten Blick widersprüchlich zu sein. Die Auflösung des dialektalen Lautsystems hat aber zur Folge, dass Dialekt mehrheitlich nicht als unabhängiges System betrachtet wird, sondern als Form der Informalität der Spontansprache, die bewusst gestaltet wird. So erscheint im intendierten Dialekt eine bewusst gestaltete Zentralisierung, welche im informelleren Gespräch spontan anders realisiert wird.

Bei den Kurzvokalen sind die Veränderungen komplexer. Einerseits werden im intendierten Dialekt die Hochzungenvokale gesenkt, nicht aber im Interview. Die Achse [i:] – [u:] bzw. [ɪ] – [ʊ] ist also relativ stabil, außer bei den Kurzvokalen im intendierten Dialekt, deren Senkung als bewusst erscheinen muss. Andererseits verändert sich die Gestaltung der Vokalräume. Die auffälligste Veränderung, welche die ganze Geometrie des Vokalraums betrifft, ist die unterschiedliche Realisierung des Kurzvokals /ɛ/. Im Vergleich

zu allen anderen Veränderungen wird es in der Spontansprache gegenüber der standardintendierten Leseaussprache nicht zentralisiert, im Gegenteil: gegenüber der Realisierung in der Lesesprache und dem intendierten Dialekt, wird /ɛ/ sogar nach vorne verlagert. Gleichzeitig wird es wie /a/ und /ɔ/ gehoben, womit es in der Spontansprache einer zentraleren Position in der Zungenhöhe entspricht. Eine solche Verteilung findet sich in der Literatur nicht. Für Halle notiert WALLRAFF (2007: 90) ein dem Standard entsprechendes /ɛ/, KEHREIN (2012: 229) und RUES u. a. (2007: 94) geben für alle e-Laute eine generelle Öffnungstendenz an, wobei die genannten Autoren nicht zwischen Leseaussprache und Interview oder weiteren Erhebungssituationen unterscheiden. Das abweichende Resultat der vorliegenden Untersuchung kann auch in diesem Fall vermutlich auf die unterschiedliche Methode zurückgeführt werden: Der ohrenphonetischen Transkription der anderen Untersuchungen steht eine instrumentalphonetische Messung gegenüber, die in dieser Beziehung genauer ist, denn die phonologische Unterscheidung, insb. zu /a/, bleibt bewahrt. Nicht gänzlich von der Hand zu weisen ist aber auch, dass eine Leipziger Besonderheit vorliegen kann, was aber ohne einen instrumentalphonetischen Vergleich mit den Hallenser bzw. Dresdener Daten nicht zu überprüfen ist. Der Artikulationsraum der Kurzvokale macht mit diesen Veränderungen für die Spontansprache einen viel ausgewogeneren Eindruck als in den beiden anderen Erhebungssituationen, insofern als Vorder- und Hinterzungenvokale nun auch wie bei den Langvokalen mehr oder weniger gleich hoch realisiert werden. Der phonetische Abstand zwischen den Vokalen bleibt damit deutlich gewahrt bzw. wird im Falle von /a/ – /ɛ/ sogar verdeutlicht. Damit tendiert das Kurzvokalsystem in der Spontansprache zu einer Optimierung der Kontraste

im eingeschränkten Raum. Es ist deshalb anzunehmen, dass dieses System relativ stabil ist.

Der allgemeine Vergleich zeigt, dass sich das Zentrum des Artikulationsraums zwischen den Erhebungssituationen verschiebt. Der über alle Sprecherinnen und alle Kardinalvokale gemittelte Wert für den ersten Formanten F_1 unterscheidet sich von Erhebungssituation zu Erhebungssituation. Während sich die Vokalhöhe im Lesetext (571 Hz) und im intendierten Ortsdialekt (563 Hz) nicht signifikant unterscheiden, ist der Wert für das Interview (545 Hz) signifikant tiefer, d. h. das Artikulationszentrum ist höher. Für diese Veränderung ist besonders die Hebung des Tiefzungenvokals /a/ bzw. /a:/ verantwortlich. Dass sich Lesetext und intendierter Ortsdialekt diesbezüglich nicht unterscheiden, ist auf den ersten Blick erstaunlich. Dies lässt sich aber einerseits damit erklären, dass die Hebung des Tiefzungenvokals im intendierten Dialekt nicht ganz so ausgeprägt ist wie im Interview, andererseits wird die Hebung durch die Senkung der kurzen Hochzungenvokale kompensiert. F_2 ist für alle drei Erhebungssituationen signifikant unterschiedlich. Am weitesten hinten realisiert ist der intendierte Dialekt (1508 Hz), etwas weiter vorne die Realisierung beim Lesen (1590 Hz) und am weitesten vorne der Wert für das Interview (1632 Hz). Die Rückverlagerung des intendierten Dialekts gegenüber der Leseaussprache ist v. a. auf die Rückverlagerung der vorderen Vokale und des Tiefzungenvokals /a/ bzw. /a:/ zurückzuführen, wohingegen die Vorverlagerung im Interview durch die Palatalisierung der hinteren Vokale zu erklären ist. Gegenüber der Leseaussprache findet sich also sowohl im intendierten Dialekt wie auch in der Spontansprache eine Verlagerung der Zungenlage. Beide gegenläufigen Veränderungen können als Zentralisierungen betrachtet werden. Diese sind aber durch Zentralisierungstendenzen in jeweils anderen Vokalgruppen

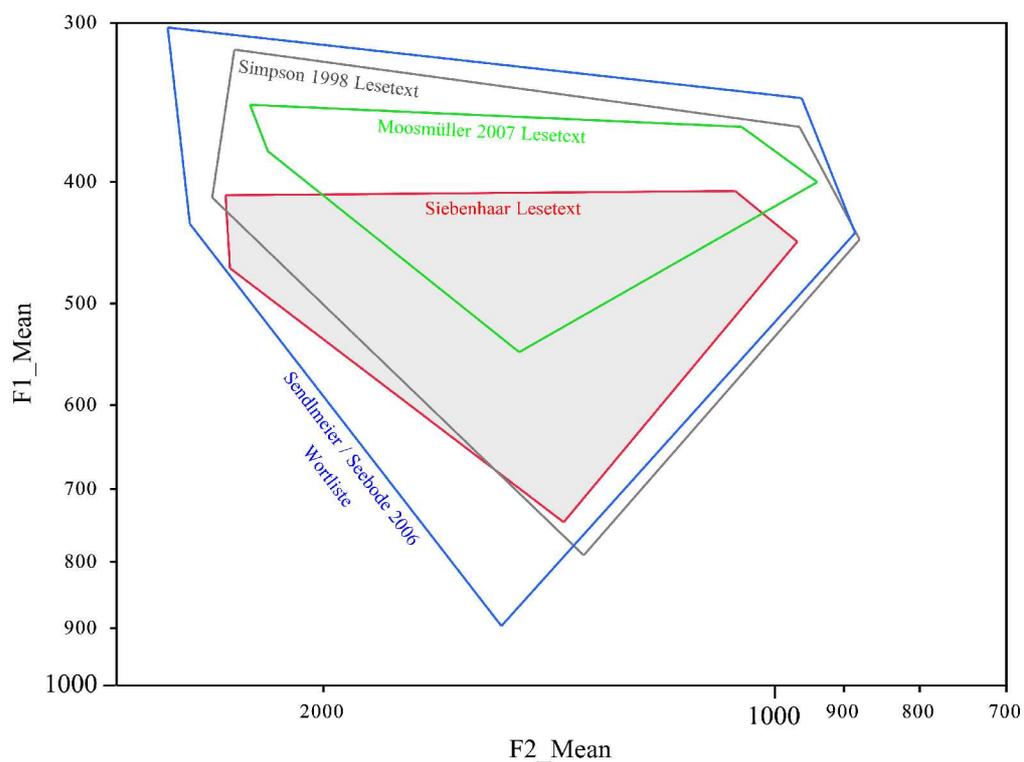
gegeben: während in der Spontansprache die Hinterzungenvokale nach vorne verlagert, also zentralisiert werden, ist im intendierten Dialekt eine Rückverlagerung, d. h. eine Zentralisierung der Vorderzungenvokale zu finden. Diese kann auf die im alten Dialekt belegte Entrundung der Vorderzungenvokale zurückgeführt werden, welche mit der Wiedereinführung dieser Laute aber nicht wieder kompensiert wird.

5.1.2 Vergleich der Messwerte der standardintendierten Leseaussprache mit anderen Untersuchungen

Ein direkter Vergleich der Messwerte mit Werten anderer Untersuchungen für die standardintendierte Lautung und eine Interpretation der vorgefundenen Unterschiede ist sehr problematisch, denn die Resultate unterscheiden sich deutlich durch die ungleichen Erhebungs- und Messmethoden, wie das hier schon als Unterschied zwischen dialektintendierten, spontansprachlichen und gelesenen Texten deutlich wird. Zudem kann keine Aufnahme als Referenzwert für die Standardlautung dienen, denn alle Aufnahmen sind jeweils zumindest auch regional geprägt. SENDLMEIER / SEEBODE (o. J. [2006]) bezeichnen ihre Aufnahmen als dialektal unmarkiert, obwohl Berliner Studierende und nicht ausgebildete Sprecher untersucht werden. Das Kiel Korpus, das von SIMPSON (1998) untersucht wird, hat eine norddeutsche Prägung, MOOSMÜLLER (2007) eine wienerische. Eine Klärung regionaler Variation im intendierten Standard wird sich vermutlich erst mit der geplanten instrumentalphonetischen Analyse des Deutsch heute-Korpus aus dem Institut für deutsche Sprache (AADG) geben. Gleichwohl soll mit aller Vorsicht ein sehr allgemeiner Vergleich mit der standardintendierten Lautung vorgenommen werden, die andernorts auch aufgenommen wurde. In den

Abbildungen 5 und 6 sind die Formantwerte der genannten Untersuchungen⁷ im Vergleich zu den eigenen Aufnahmen des Lesetextes grafisch dargestellt.

Abbildung 5: Vokalraum der Langvokale im Lesetext im Vergleich zu anderen standardintendierten Aufnahmen

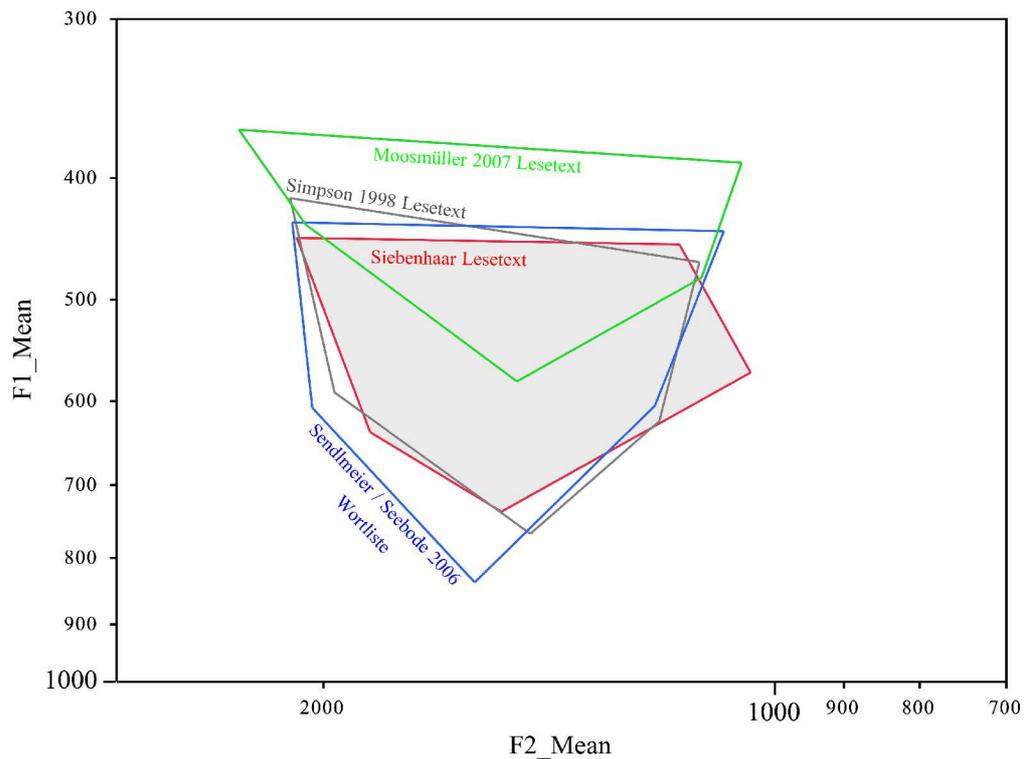


Bei den Langvokalen (Abbildung 5) ist auffällig, dass die Leipziger Sprecherinnen im intendierten Standard gegenüber den anderen Aufnahmen alle Vokale zentralisieren. Der Artikulationsraum ist dadurch deutlich verkleinert; nur die Wiener Sprecherinnen von MOOSMÜLLER (2007) zeigen

⁷ Moosmüller (2007) gibt keine Mittelwerte an. Sie führt aber die Werte der untersuchten Personen in je einzelnen Tabellen an, aus denen der Mittelwert der drei untersuchten Frauen für die vorliegende Darstellung errechnet wurde.

einen ebenso verkleinerten Artikulationsraum, dessen Zentrum jedoch viel höher liegt, was sich auch schon bei Iivonen (1994: 321) außer für [a:] gezeigt hat. Die Zentralisierung, die dem Sächsischen nachgesagt wird, zeigt sich also im Vergleich mit den anderen Aufnahmen deutlich. Allerdings ist der Vergleich mit SENDLMEIER / SEEBODE auch durch die unterschiedlichen Erhebungsmethoden beeinträchtigt, sind doch ihre Daten mit Wortlisten erhoben worden, was gegenüber Lesetexten tendenziell mit periphereren Werten einhergeht.

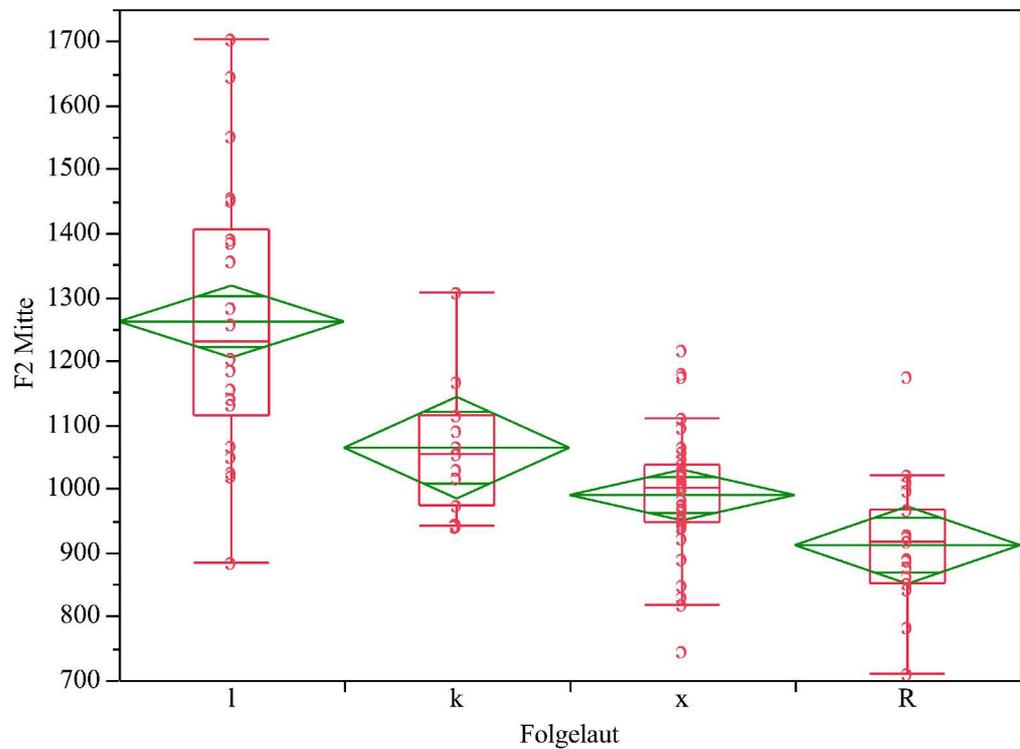
Abbildung 6: Vokalraum der Kurzvokale im Lesetext im Vergleich zu anderen standardintendierten Aufnahmen



Etwas anders sieht die Situation bei den Kurzvokalen (Abbildung 6) aus: Gegenüber den anderen Aufnahmen fällt hier insbesondere das Vokaltrapez der Wiener (MOOSMÜLLER 2007) durch die starke Hebung aller Vokale aus dem Rahmen. Von den Leipziger Aufnahmen befinden sich alle Vokale außer /ɔ/ innerhalb der Parameter von SENDLMEIER / SEEBODE (o. J. [2006]) und SIMPSON (1998). Die Zentralisierung ist aber nicht ganz so ausgeprägt wie bei den Langvokalen. Eine Sonderposition nimmt /ɔ/ ein. Der halboffene Vokal ist gegenüber den anderen Aufnahmen nach hinten verschoben. Dieser Befund widerspricht bisherigen Arbeiten, die eine Zentralisierung auditiv wahrnehmen (BECKER / BERGMANN 1969, HIRSCHFELD 1999, WALLRAFF 2007 für Halle) oder ausmessen (IIVONEN 1987, 1994). Der tiefe zweite Formant deutet auf eine große vordere Kavität und damit auf eine weit rückverlagerte Realisierung oder eine deutliche Rundung hin. Beides widerspricht der Literatur, die für das Sächsische Entrundung und Zentralisierung behauptet, wie sie sich bei den Langvokalen zeigen. Der Vergleich mit der Spontansprache (Abbildung 4) deutet auch darauf hin, dass dort ein solche Rückverlagerung nicht stattfindet. Der Blick in die Wenkersätze, welche für den intendierten Standard und auch den intendierten Dialekt erhoben wurden, macht aber deutlich, dass die Ursache für diese Datenlage in der Auswahl der Belegwörter liegt. Von den neun hier erhobenen Wörtern (*trockenen, kochen, eingebrochen, wolle, Tochter, Kochlöffel, Korb, morgen, geschmolzen*) sind nur zwei mit nicht velaren Konsonanten realisiert, sonst liegt eine Präferenz für /x/, /ʁ/ oder /k/ als Folgekonsonant vor. Diese velaren oder uvularen Konsonanten – für /ʁ/ nehmen RUES U. A. (2007) und auch KHAN / WEISE (2013) sogar eine Pharyngalisierung an – bewirken eine Rückverlagerung, und damit eine Erhöhung von F₂, die sich in den vorliegenden Daten als signifikant erweist, wie Abbildung 7 darstellt. Die Realisierung vor /l/ zeigt mit 1232 Hz

signifikant höheren Wert für F_2 , wodurch sich die Position vor nicht velaren Konsonanten im Bereich bewegt, der von Simpson (1998) oder Sendlmeier / Seebode (o.J. [2006]) angegeben wurde, womit die allgemeine Tendenz zur Zentralisierung doch auch für die Kurzvokale noch bestätigt werden kann.

Abbildung 7: F_2 -Werte von /ɔ/ nach Folgelaut, Boxplot und 95%-Konfidenzintervall innerhalb der Rhomben



5.1.3 Hinweis zu den normalisierten Werten

Die Darstellung der normalisierten Werte unterscheidet sich geringfügig von den obigen Darstellungen; der deutlichste Unterschied zeigt sich bei den ungespannten Hinterzungenvokalen [ɔ] und [ʊ], die in der normalisierten Form

etwas deutlicher voneinander getrennt sind. Zudem erscheint [ɛ] im Interview noch etwas weiter vom Zentrum weg realisiert als bei den nicht normalisierten Daten. Da die Daten mit der Normalisierung generell etwas konsistenter sind – der Grund für die Normalisierung – und weil die Werte für beiden Formanten vergleichbar sind, wird im Folgenden mit diesen normalisierten Werten weitergefahen.

5.2 Vokalraum der Individuen

Die Abbildungen 3 und 4 haben schon deutlich gemacht, dass die einzelnen Vokale je eine relativ große Streuung aufweisen. Diese Streuung ist teilweise durch die Erhebungssituationen zu erklären, wie das oben gemacht worden ist. Allerdings ist die individuelle Variation bei der Realisierung der einzelnen Vokale sehr groß. Exemplarisch zeigt das die Variation des normalisierten Wertes von F_1 des Vokals /u:/. Der Einfluss der Faktoren Person und Erhebungssituation wird mit einem Generalized Linear Model überprüft. Der Faktor Erhebungssituation (DF: 2, $F = 0.0275$, ns) erklärt nur einen sehr geringen Teil der Variation, während der überwiegende Teil durch den Faktor Person (DF: 11, $F = 4.6532$, $p < 0.001$) erklärt wird. Die individuelle Variation ist hier also bedeutend größer als die durch die Erhebungssituation. Bei anderen Vokalen ist die Erhebungssituation etwas bedeutsamer, doch generell zeigt sich eine hohe Varianz, welche mehrheitlich auf das Individuum zurückzuführen ist. Als willkürlich gewähltes Beispiel für die individuell unterschiedliche Veränderung zwischen den Erhebungssituationen wird die Lage beim Vokal /a/ dargestellt. Abbildung 8 zeigt den normalisierten F_1 -Wert, der sich umgekehrt proportional zur Vokalhöhe verhält, geordnet nach dem

Wert für die Leseaussprache. Es wird augenfällig, dass die Sprecherinnen den Wert nicht gleichmäßig ändern, sondern dass die Veränderung zwischen den Erhebungssituationen individuell unterschiedlich gestaltet wird. Im Vergleich mit der Abbildung 8, in der die entsprechenden Werte für F_2 dargestellt sind, der mit der Zungenlage einhergeht, zeigt sich auch, dass die Veränderung von F_1 (Zungenhöhe) und F_2 (Zungenlage) in keiner Weise parallel verläuft. Die Werte in Abbildung 9 sind gleich geordnet wie in Abbildung 8. Die sich daraus ergebende unorganisierte Darstellung verdeutlicht, dass sich die beiden Werte nicht gleich verhalten.

Abbildung 8: Lobanov-normalisierte F_1 -Werte von /a/ in den verschiedenen Erhebungssituationen, sortiert nach dem Wert für den Lesetext

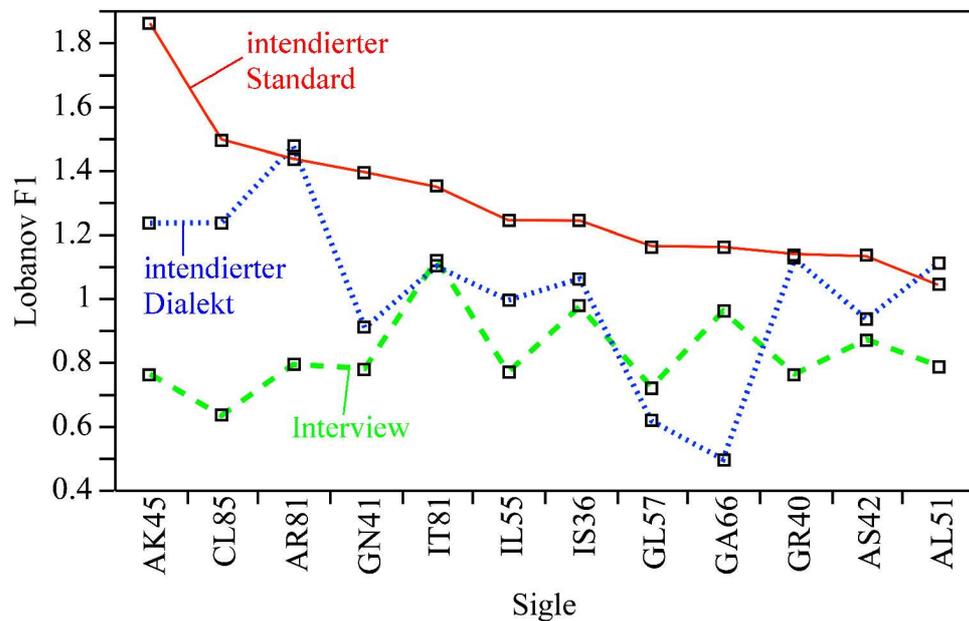
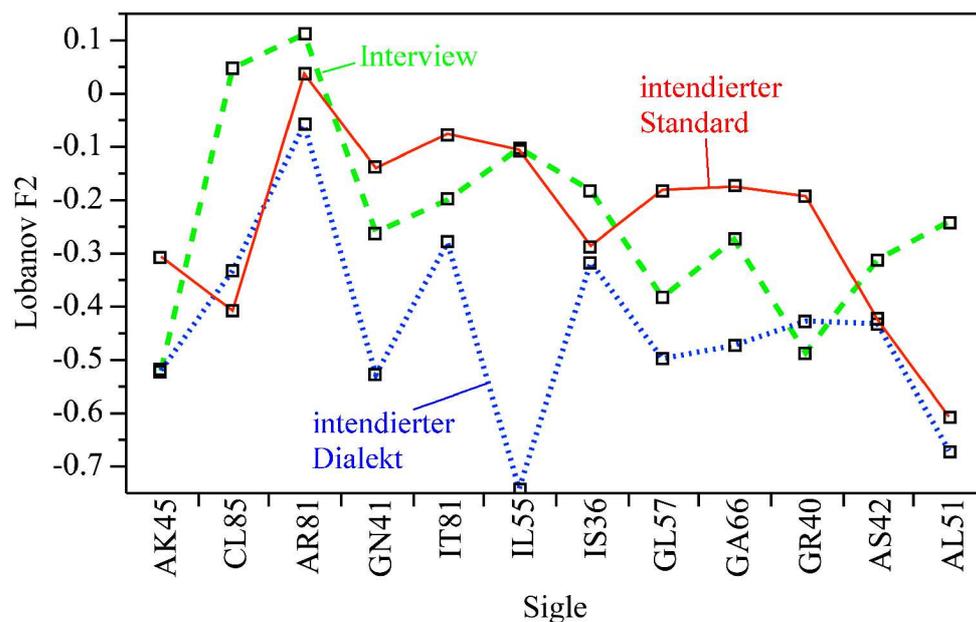


Abbildung 9: Lobanov-normalisierte F_2 -Werte von /a/ in den verschiedenen Erhebungssituationen, sortiert nach dem F_1 -Wert für den Lesetext (wie in Abbildung 8)

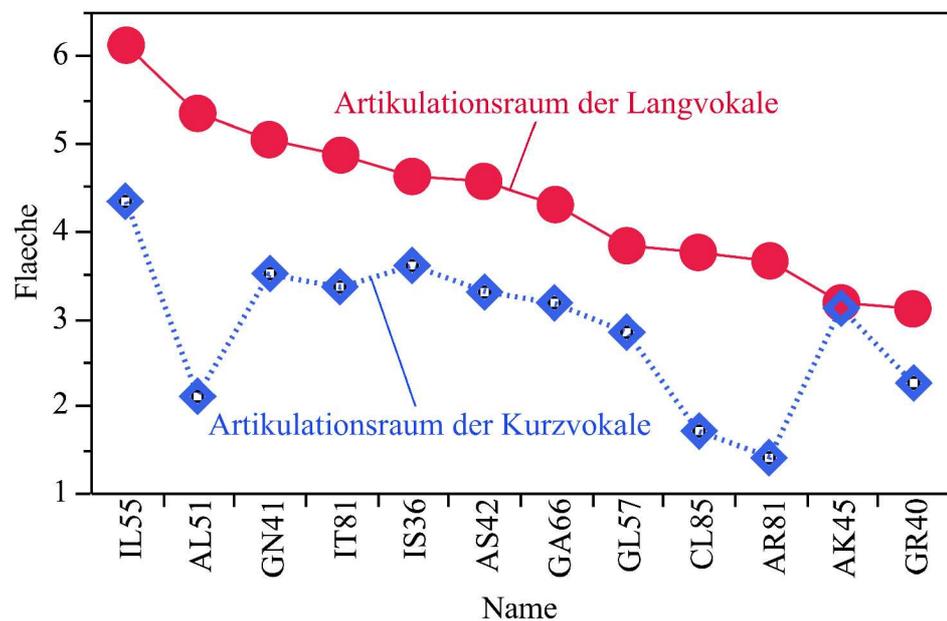


An dieser Stelle können nicht alle Vokale einzeln überprüft werden, insbesondere ging es um den Überblick und den Zusammenhang zwischen den Vokalen verloren. Deshalb wird im Folgenden der von den Individuen genutzte Vokalraum verglichen (zur Berechnung siehe Abschnitt 3.2.3). Dafür werden einerseits die Vokalräume der Individuen in der standardsprachlich intendierten Realisierung miteinander verglichen. In der Lesesituation werden mit einer Ausnahme von allen Sprecherinnen der größte Vokalraum genutzt. Damit eignet sich die Realisierung der intendierten Standardsprache als Referenzwert, mit dem die Vokalräume der anderen Erhebungssituationen verglichen werden. Andererseits werden auch die Veränderungen, welche die Individuen zwischen den Erhebungssituationen vornehmen, miteinander verglichen. Um diese

individuellen Unterschiede darzustellen, wird für diese Analysen auf die normalisierten Daten zurückgegriffen und daraus die Fläche berechnet.

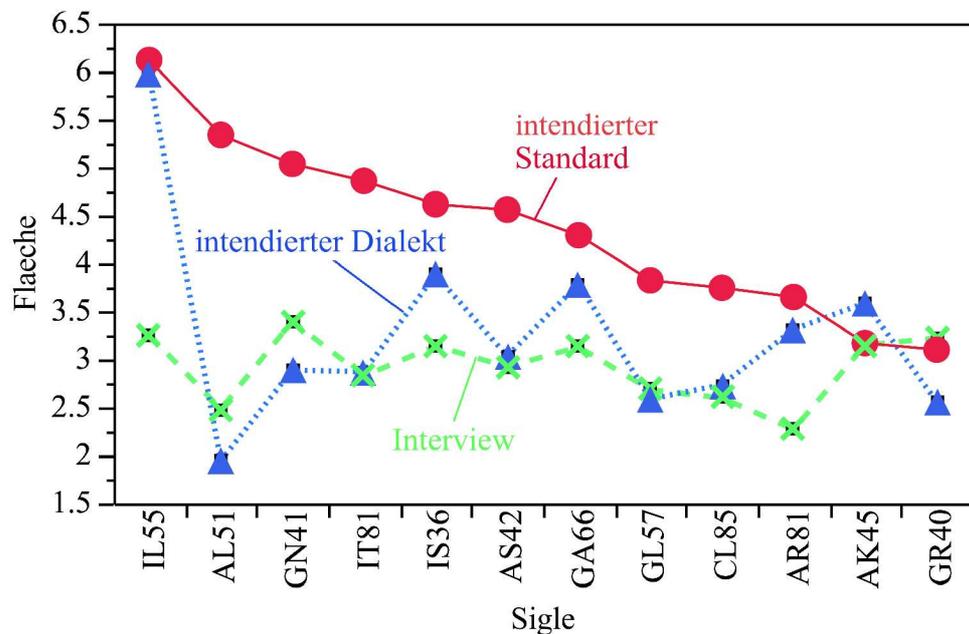
Abbildung 10 zeigt die Größe des genutzten Vokalraums in der standardintendierten Lesesituation für Langvokale und Kurzvokale. Dabei ist evident, dass die Kurzvokale bei allen Sprecherinnen einen kleineren Raum einnehmen als die Langvokale, allerdings ist die Reduktion des Artikulationsraumes je nach Sprecherin unterschiedlich groß. Auffällig ist der sehr geringe Unterschied der Artikulationsraumgröße zwischen Lang- und Kurzvokalen bei AK45 und der sehr große Unterschied bei AL51.

Abbildung 10: Größe des Artikulationsraums in der standardintendierten Lesesituation in z-normalisierten quadrierten Einheiten nach Langvokalen und Kurzvokalen



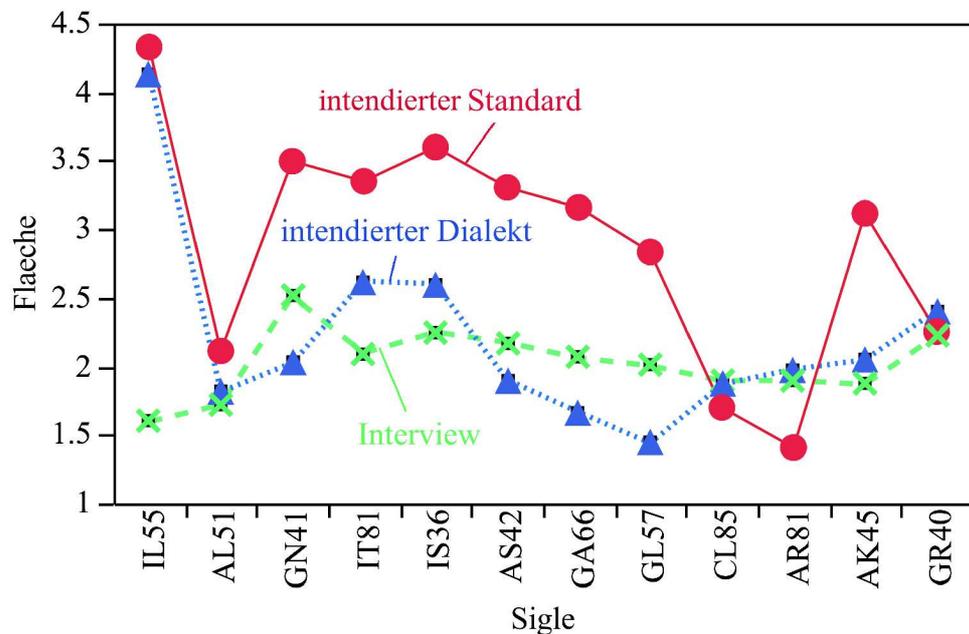
Die Abbildungen 11 und 12 zeigen die unterschiedliche Größe der Vokalräume für Lang- bzw. Kurzvokale in Abhängigkeit von der Erhebungssituation. Bei den Langvokalen (Abbildung 11) waren die meisten Sprecherinnen den Unterschied zwischen der intendierten Standardsprache und den beiden anderen Erhebungssituationen, die meist einen kleineren Artikulationsraum einnehmen. Einzig AK45 zeigt im intendierten Dialekt einen größeren Artikulationsraum als im intendierten Standard. Zwei Drittel der Sprecherinnen zeigen im intendierten Dialekt einen größeren Artikulationsraum als im Interview, was evtl. durch die langsamere Artikulation im bewussten Sprechvorgang erklärt werden kann. Auffällig ist in dieser Hinsicht besonders die Sprecherin IL55, die für den intendierten Dialekt einen fast ebenso großen Artikulationsraum nutzt wie für das standardintendierte Vorlesen.

Abbildung 11: Größe des Artikulationsraums der Langvokale in z-normalisierten quadrierten Einheiten nach Erhebungssituationen



In Abbildung 12 mit dem Artikulationsraum der Kurzvokale sind die Sprecherinnen wiederum in der selben Reihenfolge wie in Abbildung 11 aufgeführt, deshalb die unruhige Darstellung. Für die meisten Sprecherinnen ist der Artikulationsraum in der intendierten Standardausprache auch bei den Kurzvokalen größer als in den beiden anderen Erhebungssituationen. Jedoch sind bei dreien der Artikulationsraum im intendierten Dialekt und bei zweien auch derjenige des Interviews größer. Auffällig ist wiederum IL55, deren Artikulationsraum auch bei den Kurzvokalen im intendierten Dialekt fast ebenso groß ist wie im intendierten Standard. Bei den meisten anderen Sprecherinnen ist der Unterschied zwischen intendiertem Dialekt und Interview kleiner als bei den Langvokalen.

Abbildung 12: Größe des Artikulationsraums der Kurzvokale in z-normalisierten quadrierten Einheiten nach Erhebungssituation



Insgesamt wird also der Artikulationsraum besonders zwischen der Leseaussprache und den Realisierungen in den beiden anderen Erhebungssituationen verkleinert, die sich etwas weniger voneinander unterscheiden. Deutlich abweichend verhält sich die Sprecherin IL55, die für den intendierten Standard und den intendierten Dialekt sowohl bei Langvokalen wie auch bei den Kurzvokalen einen je fast gleich großen Artikulationsraum aufweist, den sie überaus klar vom deutlich verkleinerten Artikulationsraum im Interview abhebt. Auf der anderen Seite des Variationsspektrums ist die Sprecherin GR40, die in allen drei Erhebungssituationen nur sehr geringe Unterschiede realisiert. AL51 zeigt immer sehr kleine Artikulationsräume außer bei den Langvokalen der Leseaussprache, wo sie den zweitgrößten Artikulationsraum aller Gewährsfrauen aufweist.

5.3 Lautdauer und Artikulationsraum

An dieser Stelle wird nochmals der Zusammenhang zwischen Lautdauer und Artikulationsgenauigkeit aufgegriffen, um die Undershoot-These, dass kürzere Laute zur Zentralisierung tendieren, zu überprüfen. Als Korrelat der Artikulationsgenauigkeit wird hier die Größe des Artikulationsraumes angesetzt. Abbildung 13 zeigt zur Einbettung der folgenden Analysen die durchschnittliche Länge der Lang- bzw. Kurzvokale nach den Erhebungssituationen.⁸ Der Unterschied zwischen Lang- und Kurzvokalen

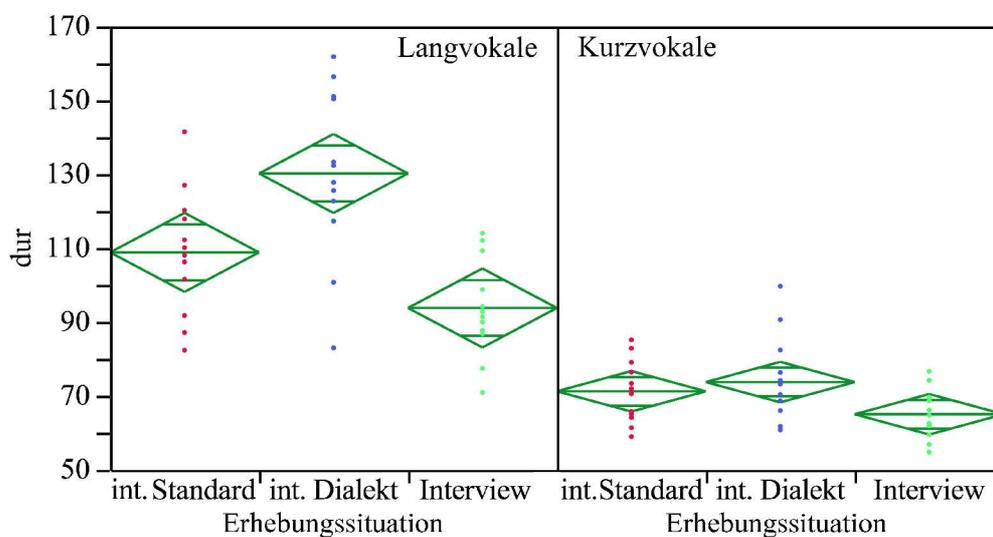
⁸ Die Datenmenge ist zu klein um die Vokale einzeln zu berücksichtigen. Diese Differenzierung wird häufig vorgenommen, da Tiefzungenvokale im allgemeinen länger sind

bleibt in allen Erhebungssituationen gewahrt. Er ist im intendierten Dialekt am größten, wo die Kurzvokale 57% der Länge der Langvokale aufweisen. In der Leseaussprache entspricht die Länge der Kurzvokale noch 65% der Länge der Langvokale, und im Interview reduziert sich der Unterschied weiter, wo die Kurzvokale schon 69% der Länge der Langvokale ausmachen. Damit liegt die Verteilung für die Leseaussprache leicht außerhalb dessen, was bisherige Untersuchungen festgehalten haben: RAMERS (1988: 197–200) hat für die Standardsprache, je nach Sprecher, ein Verhältnis von Lang- zu Kurzvokalen zwischen 1 : 0.39 und 1 : 0.61 erfasst. SPIEKERMANN (2000: 221–222) belegt für den Standard 1 : 0.60. Der hier belegte Wert von 1 : 0.65 für die Standardsprache zeigt eine leicht weniger ausgeprägte Trennung von Lang- und Kurzvokalen. Für Dialekte belegt SPIEKERMANN Verhältnisse zwischen 1 : 0.39 (Rheinfränkisch) und 1 : 0.82 (Mittelbayrisch). Für Wurzen 30 km östlich von Leipzig ist ein Verhältnis von 1 : 0.66 belegt. Die etwas stärkere Differenzierung in den vorliegenden Daten (1 : 0.57) kann individuell begründet sein – SPIEKERMANN hat nur eine einzige Aufnahme untersucht und bei den vorliegenden zwölf Aufnahmen bewegt sich der Wert für den intendierten Dialekt auch zwischen 1 : 0.44 und 1 : 0.74. Nicht von der Hand zu weisen ist auch ein Einfluss der Erhebungssituation, denn beim intendierten Dialekt muss öfters von einer Stilisierung ausgegangen werden. Diese Stilisierung manifestiert sich oft besonders deutlich in der Prosodie, und damit in der Dauer, weil die starke Betonung akzentuierter Silben die Opposition zwischen Lang- und Kurzvokalen verstärkt .

als Hochzungenvokale. Diese Tatsache ist von MADDIESON 1997 sogar als phonetische Universale bezeichnet wurde, was in einzelnen Untersuchungen (MOOSMÜLLER 2007: 71 für das österreichische Deutsch) aber zumindest angezweifelt wurde.

Während die Erhebungssituationen bei den Langvokalen hoch signifikant unterschiedliche Längen zeigen (ANOVA, $F = 12.11$, $DF = 2$, $p = 0.0001$), ist der Unterschied bei derselben Tendenz bei den Kurzvokalen nicht signifikant. Bei Kurz- und Langvokalen sind die Vokale im intendierten Dialekt länger als in der Lesesprache, was vermutlich mit der erhöhten Aufmerksamkeit durch die Übersetzungsleistung, wie auch mit der bewusst genaueren Artikulation in der Präsentationssituation zu tun hat. Die kürzere Dauer im Interview ist dagegen wohl auf die allgemein erhöhte Sprechgeschwindigkeit in der Spontansprache und die in dieser Situation weniger bewusst gestaltete Realisierung zurückzuführen.

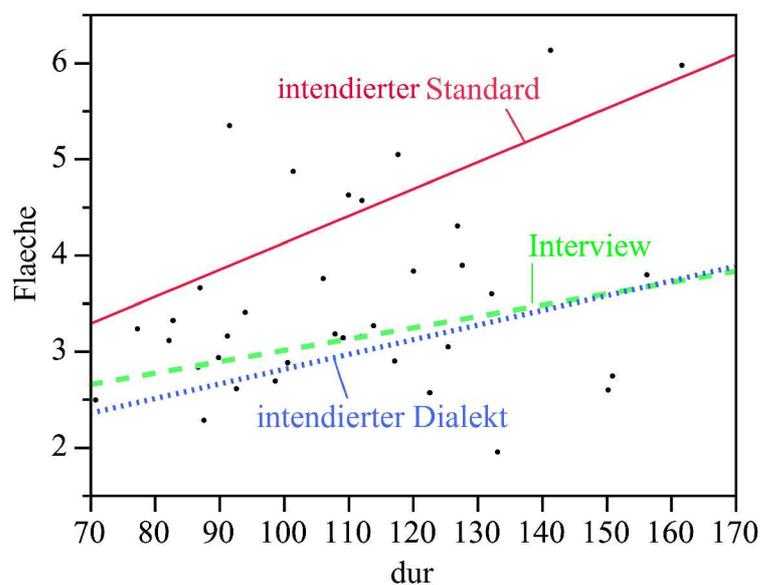
Abbildung 13: durchschnittliche Länge der Langvokale / Kurzvokale nach Erhebungssituation. Die Rhomben bezeichnen das 95%-Konfidenzintervall



Eine Kürzung der Lautdauer wird auch mit der Reduktion der Artikulationsgenauigkeit in Verbindung gebracht. Es wird deshalb für jede Erhebungssituation geprüft, ob die individuelle Veränderung der Lautdauer mit

der Größe des Artikulationsraumes korreliert. Die Abbildungen 14 und 15 zeigen die Korrelation zwischen Vokallänge und Größe des Artikulationsraumes für jede Sprecherin. Wiederum werden für die Berechnung des Artikulationsraums die z-normalisierten Werte verwendet. Für die Langvokale (Abbildung 14) wird für alle Erhebungssituationen ein leichter Zusammenhang von Größe des Artikulationsraumes und der Vokallänge ersichtlich, freilich ist dieser in keinem der Fälle auf dem 5%-Niveau signifikant; es liegen allerdings für jede Erhebungssituation auch nur zwölf Werte – für jede Person einen – vor.

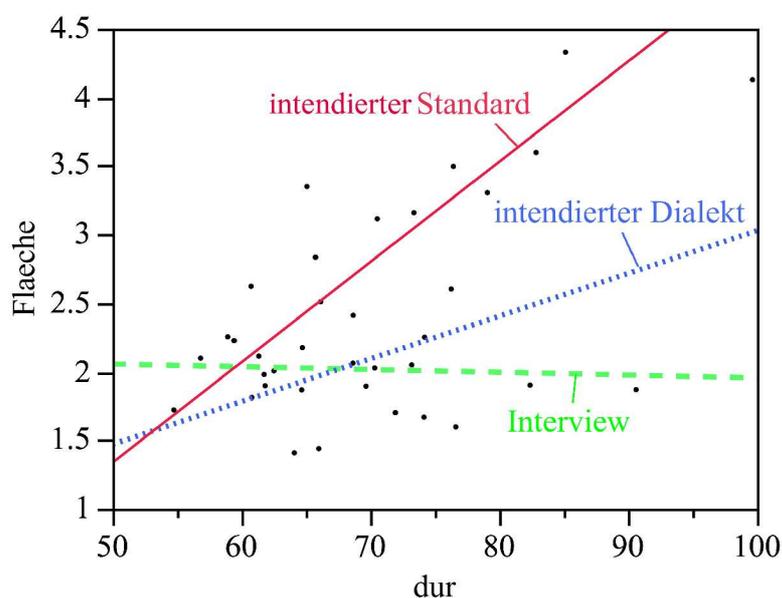
Abbildung 14: Zusammenhang von Lautdauer der Langvokale in ms und Artikulationsraum in z-normalisierten quadrierten Einheiten nach Erhebungssituationen



Ganz anders sieht die Situation für die Kurzvokale aus (Abbildung 15). Während nämlich sowohl für die Standardrealisierung als auch für den intendierten Dialekt ein signifikanter positiver Zusammenhang zwischen

Vokallänge und Größe des Artikulationsraumes besteht, der für den intendierten Standard viel ausgeprägter ist als für den intendierten Dialekt, ist der Zusammenhang in der Interviewsituation nicht vorhanden.

Abbildung 15: Zusammenhang von Lautdauer der Kurzvokale in ms und Artikulationsraum in z-normalisierten quadrierten Einheiten nach Erhebungssituationen



Damit ist klar, dass die Verkleinerung des Artikulationsraumes in der Spontansprache nicht mit Undershoot, dem Nicht-Erreichen der artikulatorischen Zielpunkte, erklärt werden kann. Vielmehr kann die relativ geringe Größe des Artikulationsraumes bei den Kurzvokalen selbst als Merkmal der Leipziger Spontansprache gelten. Der Vergleich der beiden Abbildungen deutet auf einen artikulatorischen Aspekt hin, der vermutlich für die Perception des Sächsischen relativ bedeutsam ist. Das von Laien für das Sächsische oft als charakteristisch genannte Nuscheln, was sich ja v. a. auf die

im Alltag wahrgenommene Spontansprache mit regionaler Färbung bezieht, findet an dieser Stelle seine artikulatorische Erklärung. Zuallererst ist die kürzere Realisierung aller Vokale in der Spontansprache zu erwähnen, die wohl nicht spezifisch für das Sächsische ist, aber zum allgemeinen Eindruck beiträgt. In regionaler Hinsicht bedeutsam sind aber zwei andere Faktoren: Einerseits wird die phonologische Opposition zwischen Lang- und Kurzvokalen in der Spontansprache weniger deutlich durch Dauerunterschiede markiert als in der Leseaussprache. Andererseits werden die einzelnen Vokale auch bei stärkerer Dehnung weniger deutlich artikuliert, als das im intendierten Standard der Fall ist, für die Kurzvokale ist sogar eine schwache Tendenz zu einer weniger deutlichen Realisierung, d. h. einer Zentralisierung unter Dehnung festzustellen.

5.4 Typen von Sprecherinnen

Die vorigen Kapitel haben deutlich gemacht, dass die Sprecherinnen den Artikulationsraum von der Lesesituation über die intendierte Mundart zur Interviewsituation verkleinern, dass diese Veränderung aber von Sprecherin zu Sprecherin unterschiedlich ausgeprägt ist. Ein Zusammenhang mit der Kürzung der Vokale konnte nur sehr eingeschränkt hergestellt werden. Es kann also vermutet werden, dass sich die Sprecherinnen in ihren Veränderungen individuell verhalten. Möglich ist aber auch, dass mehrere dieser Sprecherinnen ähnliche Veränderungsmuster zeigen und sich die Veränderungen somit nicht als individuell sondern als typisch erweisen. Um das zu überprüfen, wird mit Methoden der explorativen Modellierung nach Sprecherinnen mit ähnlichen Veränderungsmustern gesucht. Dafür wird häufig

ein so genanntes Clustering (SAS 2009: 560–563; BAAYEN 2008: 138–154) verwendet, eine Gruppe statistischer Verfahren, die für Markt- und Wählerforschung etabliert sind und auch schon in variationslinguistischen Arbeiten genutzt wurden (LENZ 2003: 217–245; LAMELI 2004: 193–203). Mittels Clustering werden Objekte zusammen gruppiert, welche in verschiedenen Variablen jeweils ähnliche Variantenausprägungen aufweisen. Die hier verwendete hierarchischen Clusteranalyse fasst jeweils die zwei Sprecherinnen mit der geringsten Distanzrelation zu einer Gruppe, einem Cluster, zusammen. Das Verfahren wird so lange mit einzelnen Sprecherinnen bzw. Clustern weitergeführt, bis alle Elemente in einem Cluster zusammenfallen. Für die Interpretation wird dann eine als sinnvoll erachtete Anzahl Cluster bestimmt – eine feste Größe gibt es nicht –, welche in ihren je gemeinsamen Merkmalen interpretiert werden. Die Darstellung der Gruppierung erfolgt in einem Dendrogramm.

Für die Gruppierung kann eine nur durch die Rechenkapazität begrenzte Anzahl Variablen berücksichtigt werden. Im Rahmen dieser Analyse wurden verschiedene Kombinationen durchgespielt. Genauer gezeigt werden hier zwei Analysen. Die eine berücksichtigt die Größe der Vokalräume, die zweite die Vokaldauer, wobei für jede der beiden Analysen Lang- und Kurzvokale in den verschiedenen Erhebungssituationen, also je lediglich sechs Variablen, in die Berechnung einfließen. Bemerkenswert ist, dass sich die Berechnung mit der Größe der Vokalräume als relativ stabil erwiesen hat, denn die Personen werden genau gleich gruppiert, wie wenn die Formantwerte aller Vokale mitberücksichtigt werden, was rund 60 Variablen entspricht. Eine solche Darstellung ist aber insgesamt schwieriger zu interpretieren. Für artikulatorische Veränderungen der Vokale ist damit die Größe der Vokalräume ein guter Indikator. Eine weitere Clusteranalyse, welche

gleichzeitig die Dauer als auch die Qualität berücksichtigt, wird zum Schluss des Abschnitts diskutiert.

Die Abbildung 16 stellt dieses Dendrogramm mit der Größe der Vokalräume als Gliederungskriterien dar. Als erstes wird die Sprecherin IL55 von den anderen abgesondert. Deren Eigenständigkeit hat sich schon im Vergleich der Artikulationsräume (vgl. Abbildung 11 und Abbildung 12) abgezeichnet, insbesondere durch den gegenüber dem intendierten Standard kaum reduzierten Artikulationsraum im intendierten Dialekt. Sie zeigt in allen Fällen den größten Artikulationsraum aller Sprecherinnen außer bei den Kurzvokalen im Interview, wo sie sogar den kleinsten Artikulationsraum aufweist. Die übrigen Sprecherinnen lassen sich dann in zwei Gruppen unterscheiden. Der Cluster mit AR81, CL85, GL57 und AL51 zeichnet sich v. a. dadurch aus, dass der Unterschied zwischen intendiertem Dialekt und Interview eher klein ist und sich gleichzeitig Lang- und Kurzvokale wenig unterscheiden. Die vier Sprecherinnen zeigen den kleinsten Artikulationsraum der Langvokale im Interview. Der dritte Cluster unterscheidet sich vom zweiten dadurch, dass die Sprecherinnen im Interview deutlich größere Artikulationsräume nutzen. Zudem ist auch der Artikulationsraum der Kurzvokale in der Leseaussprache größer. Innerhalb dieser Gruppe können noch AK45 und GR40 abgesondert werden: Gegenüber den anderen Sprecherinnen dieses mittleren Clusters verändern sie den Artikulationsraum für Langvokale in der Leseaussprache gegenüber den anderen Erhebungssituationen nicht.

Abbildung 16: Dendrogramm mit Größe des Artikulationsraumes nach Kurz-, Langvokal und Erhebungssituation

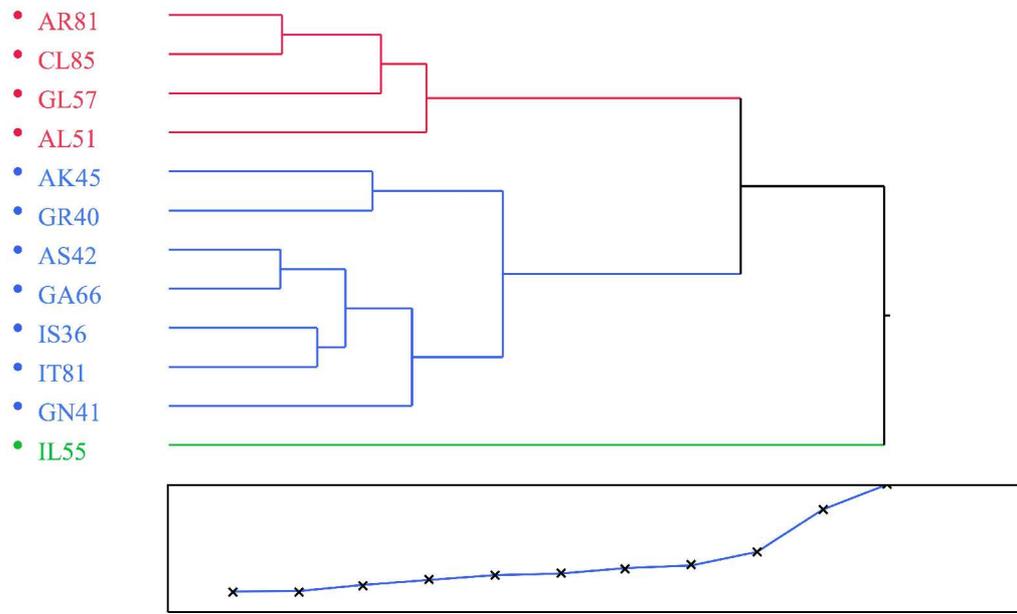
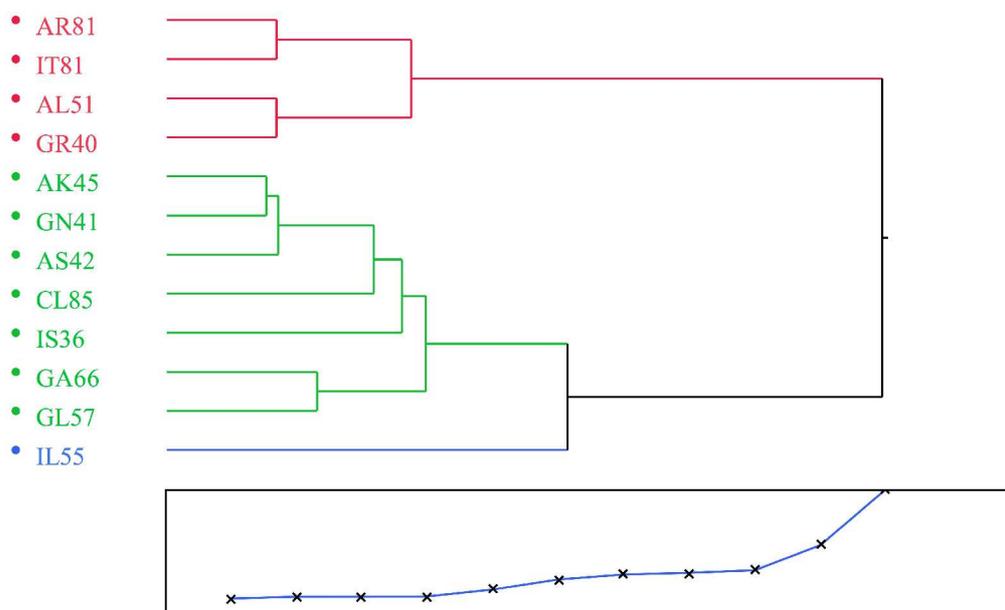


Abbildung 17 zeigt ein Dendrogramm, das die Vokallänge der Lang- und Kurzvokale in den verschiedenen Erhebungssituationen als Input berücksichtigt. Der Cluster AR81, IT81, AL51 und GR40 realisiert beide Vokalklassen in allen Situationen am schnellsten. Von der anderen, langsameren Gruppe trennt sich IL55 mit der jeweils längsten Dauer aller Vokale.

Abbildung 17: Dendrogramm mit Vokaldauer nach Kurz-, Langvokal und Erhebungssituation



Der Vergleich der beiden Dendrogramme weist darauf hin, dass Vokaldauer und Größe des Vokalraumes nicht gleichartig verändert werden. Einzig die Sonderstellung von IL55 in beiden Clusteranalysen ist auffällig. Eine genauere Analyse erlaubt die Verteilung innerhalb des kleinen Samples leider nicht. Auch weitere Interpretationen sind durch die geringe Datenmenge nur als vage Tendenzen aufzufassen. Ein Zusammenhang mit dem Alter der Gewährsfrauen – ihr Jahrgang ist jeweils in der Sigle enthalten – ist sichtbar. Der jeweils oben dargestellte Cluster mit vier Mitgliedern ist je etwas jünger als der andere, der Altersunterschied ist für den Vokalraum signifikant (ANOVA, $F = 11.64$, $DF = 2$, $p = 0.0001$), nicht jedoch für die Lautdauer. Ob dieser Unterschied sich aber als stabil erweist und damit als Hinweis für eine systematische

Altersdifferenzierung im Bezug auf die Artikulation anzusehen ist, ist erst an einem größeren Datensatz zu überprüfen.

Eine weitere Clusteranalyse ist durchgeführt worden, welche gleichzeitig die Qualität als auch die Dauer berücksichtigt. Diese Analyse zeigt für eine Aufteilung in drei Cluster, die wiederum als relativ klare Strukturierung erscheint, eine Struktur, die der Analyse mit der Dauer entspricht. Die Sonderstellung von IL55 wird aber noch etwas deutlicher. Die Sprecherinnen sind also im Bezug auf die qualitativen Veränderungen ähnlicher als im Bezug auf die Daueränderungen. Das ist insofern erwartbar, als die Dauer phonologisch nur zwei Kategorien aufweist, die von den Individuen auch jeweils auseinandergehalten werden. Der Variationsraum ist dadurch generell größer, was individuell ausgestaltet werden kann, ohne dass das Lautsystem tangiert wird. Allerdings ist noch nicht abzuschätzen inwiefern dieses Clustering auf die jeweiligen Messverfahren zurückzuführen ist.

6 Zusammenfassung und Interpretation

Die unterschiedlichen Aufnahmesituationen – Interview, Lesetext und intendierter Dialekt – zeigten deutlich unterschiedlich strukturierte Artikulationsräume. Leipziger unterscheiden also ihre Vokale in verschiedenen Sprechsituationen.

Die Änderung der Lautqualität in den dialektintendierten Wenkersätzen und in der Spontansprache gegenüber der Leseaussprache zeigen je Tendenzen zur Zentralisierung, allerdings werden je andere Vokalgruppen zentralisiert: Während in der Spontansprache die Hinterzungenvokale nach vorne verlagert, also zentralisiert werden, ist im intendierten Dialekt eher eine

Rückverlagerung, d. h. eine Zentralisierung der kurzen Vorderzungenvokale zu finden. Im intendierten Dialekt werden zudem die kurzen Hochzungenvokale gesenkt. [i:] erscheint über alle Erhebungssituationen hinaus als Fixpunkt. Auch die Lautdauer wird zwischen den Erhebungssituationen unterschieden; dabei ist die Vokaldauer in der Spontansprache am kürzesten, am längsten ist sie im intendierten Dialekt. Somit lässt sich die Zentralisierung im intendierten Dialekt nicht allein mit der Kürzung als Ursache für einen artikulatorischen Undershoot erklären, sondern weist auf andere Artikulationsintentionen hin, insbesondere auf eine bewussteren Artikulation in der Lesesprache und noch stärker im intendierten Dialekt. Diese erhöhte Bewusstheit für die Artikulation äußert sich auch in der akustischen Realisierung der Längenopposition, die im intendierten Dialekt am deutlichsten ist, gefolgt von der Opposition im Lesetext und dem geringsten Unterschied in der Spontansprache. Somit ist die Spontansprache artikulatorisch nicht einfach eine Zwischenform von Dialekt und Standard. Sie ist auch nicht die Aussprache des Hochdeutschen in der Lautung des Dialekts, wie das BECKER / BERGMANN (1969: 142) in der lautlichen Beschreibung der obersächsischen Umgangssprache erklärt haben, wobei der dort grammatisch-strukturell zu verstehende Begriff Umgangssprache trotz gewisser Überschneidungen nicht ohne Weiteres mit dem funktional-stilistischen Begriff der Spontansprache gleichgesetzt werden darf. Die artikulatorische Sonderstellung der Spontansprache hängt eindeutig damit zusammen, dass diese generell weniger genau artikuliert wird als die bewussteren Sprechweisen wie in der vorliegenden Untersuchung die gepflegtere standardintendierte Leseaussprache oder die im Alltag kaum mehr verwendete dialektalere Sprechlage. Die geringere Artikulationsgenauigkeit der Spontansprache manifestiert sich in Kürzungen und Zentralisierungen, welche letztere tendenziell eher auf das System des intendierten Standards zu

beziehen sind als auf das mundartlich intendierte System. Deutlich ist das besonders an der Qualität der A-Laute. Diese sind in der Spontansprache gegenüber der standardintendierten Lautung jeweils gehoben, die mundartlich intendierte Lautung ist dagegen weiter hinten realisiert und unterscheidet sich damit in der Zungenlage deutlich von der standardintendierten und der spontansprachlichen Lautung. Die mundartlich intendierte Lautung zeigt gegenüber der standardintendierten auch Zentralisierungstendenzen der Vorderzungenvokale, was auf die ursprüngliche Entrundung der sächsischen Mundarten zurückzuführen ist, welche aber in den aktuellen Aufnahmen nur noch vereinzelt vorkommt. Ebenfalls sind die kurzen Hochzungenvokale in der intendierten Mundart deutlich gesenkt, was in der Spontansprache nicht vorkommt. Das Backing der Vorderzungenvokale und der A-Laute sowie die Senkung der kurzen Hochzungenvokale sind damit Merkmale, die bewusst eingesetzt werden, um Mundart zu artikulieren. Zudem sind auch die langen Hinterzungenvokale in der intendierten Mundart gegenüber der Leseaussprache etwas zentralisiert, allerdings nicht in dem Maße wie in der Spontansprache. Das kann ebenfalls als bewusste Gestaltung beurteilt werden, weil diese Zentralisierung nicht mit einer Kürzung der Vokaldauer einhergeht, wie das in der Spontansprache geschieht.

Die Analysen der einzelnen Vokalqualitäten haben einerseits Unterschiede zwischen den Gewährspersonen innerhalb derselben Erhebungssituationen gezeigt. Andererseits wurde auch deutlich, dass Sprecherinnen die Qualität der Vokale zwischen den Erhebungssituationen je anders verändern. Um diese Veränderungen strukturiert zu beobachten, wurde die Größe des Artikulationsraumes verglichen. Insgesamt wird der Artikulationsraum besonders zwischen der standardintendierten Leseaussprache und den beiden anderen Erhebungssituationen verkleinert, die sich in der Größe der

Artikulationsräume etwas weniger voneinander unterscheiden. Die Ausgestaltung der Veränderungen des Vokalraumes ist allerdings individuell überaus unterschiedlich. Dadurch sind die oben genannten allgemeinen Artikulationstendenzen durch individuelle Variation zumindest teilweise überlagert. Eine Clusteranalyse, welche die Ähnlichkeit zwischen Personen als Kategorisierungskriterien zugrunde legt, hat eine einzelne Sprecherin von allen übrigen abgegrenzt. Diese weist auffällig geringe Unterschiede zwischen dem intendierten Standard und dem intendiertem Dialekt auf und artikuliert in diesen beiden Erhebungssituationen sehr genau. Die übrigen Sprecherinnen unterscheiden sich durch unterschiedliche Reduktion des Artikulationsraumes zwischen intendiertem Dialekt und Interview, wobei diejenigen, welche intendierten Dialekt und Spontansprache weniger unterscheiden, etwas jünger sind als diejenigen mit größeren Unterschieden. Das deutet darauf hin, dass die jüngeren die intendierte Mundart weniger genau als eigene Sprechlage gestalten als die älteren. Für eine präzisere Aussage müssten dafür aber einerseits mehr Gewährspersonen zur Verfügung stehen, andererseits müssten dann auch wieder die Einzellaute berücksichtigt werden.

Die Arbeit differenziert damit die für Dresden (KEHREIN 2012) festgestellte geringe Unterscheidung verschiedener Sprechlagen, wobei hier offen bleiben muss, ob die festgestellten Unterschiede zwischen den beiden sächsischen Städten als regional bedingt zu interpretieren sind oder ob dies an den unterschiedlichen Methoden liegt. Das hier verwendete instrumentalphonetische Messverfahren kann jedoch KEHREINS Ergebnis bestätigen, dass Zentralisierung einen wesentlichen Aspekt im Sächsischen darstellt. Es zeigt aber auch, dass die Zentralisierung je nach Sprechlage unterschiedlich realisiert wird, womit die bisherigen ohrenphonetischen Arbeiten um einen wesentlichen Punkt ergänzt werden. Sämtliche Analysen

haben aber auch deutlich gemacht, dass die individuellen Unterschiede beträchtlich sind. Die Clusteranalyse kann einzelne Sprechertypen mit gemeinsamen Variationsstrukturen herausarbeiten. Die Datenmenge ist aber zu klein, als dass die vorliegenden Resultate eine allgemeine Gültigkeit für sich beanspruchen könnten. Die typische Sächsischsprecherin konnte damit leider noch nicht erfasst werden, denn dafür wäre auch eine Berücksichtigung weiterer Regionen in Sachsen und außerhalb notwendig. Für den Vergleich des intendierten Standards wären auch professionelle Sprecherinnen zu berücksichtigen, welche als überregionale Referenz gelten könnten.

In methodischer Hinsicht erweist sich die als Datenreduktionsverfahren verwendete Berücksichtigung der Größe des Artikulationsraumes statt der Menge aller Vokale als stabil und für die verwendeten Vergleiche als ausreichend. Der Verlust der genauen Vokalpositionen ist aber mit zusätzlichen Messpunkten, z. B. dem hier versuchsweise verwendeten Zentralwert aufzufangen, wobei allenfalls andere Möglichkeiten erarbeitet werden sollen. Instrumentalphonetische Zugänge stellen damit einmal mehr eine sinnvolle Ergänzung klassisch dialektologischer Arbeitsweisen dar, wobei in den hier schon aufbereiteten Daten noch Potenzial liegt, da im Rahmen der Studie viele Parameter vereinfacht worden sind, welche weitere Differenzierungen ermöglichen. Zu erwähnen ist dabei besonders die Einschränkung auf einen einzigen Messpunkt innerhalb der Vokale, wodurch der häufig leicht diphthongische Charakter der meisten Vokale nicht erfasst worden ist. Eine solche Änderung ermöglichte zudem den Einbezug der Diphthonge. Im Weiteren wäre die lautlich Umgebung der einzelnen Vokale, wie es in der klassischen Dialektologie Standard ist und wie es hier in der Analyse der F_2 -Werte vor [ɔ] exemplarisch gezeigt wurde, systematisch zu erfassen, was aufgrund der teilweise noch geringen Datenmenge hier nicht zu leisten war.

Zum Schluss folgt deshalb noch das *Ceterum censeo* der meisten empirischen linguistischen Arbeiten: Gesicherte Ergebnisse sind nur mit mehr Daten zu erreichen.

Literaturverzeichnis

- AADG = Atlas zur Aussprache des deutschen Gebrauchsstandards. In: <http://prowiki.ids-mannheim.de/bin/view/AADG/WebHome> (12.03.2014).
- ALBRECHT, KARL (1881 (Nachdruck: 1983)): Die Leipziger Mundart: Grammatik und Wörterbuch der Leipziger Volkssprache; zugleich ein Beitrag zur Schilderung der Volkssprache im allgemeinen. Frankfurt a. M.: Weidlich.
- ANDERS, CHRISTINA ADA (2010): Wahrnehmungsdialektologie: das Obersächsische im Alltagsverständnis von Laien. Berlin u. a.: de Gruyter. (= Linguistik Impulse & Tendenzen 36)
- ANDROUTSOPOULOS, JANNIS (2012): "Intermediale Varietätendynamik: Ein explorativer Blick auf die Inszenierung und Aushandlung von ‚Dialekt‘ auf YouTube". *Sociolinguistica* 26: 87–101.
- BAAAYEN, ROLF HARALD (2008): *Analyzing Linguistic Data*. Cambridge: CUP.
- BARTHEL, FRIEDRICH (1933): Der vogtländisch-westerzgebirgische Sprachraum. *Kulturgeographische Untersuchungen zum Grenzproblem*. Halle an der Saale. (= Mitteldeutsche Studien 5)
- BAUMGÄRTNER, KLAUS (1959): *Zur Syntax der Umgangssprache in Leipzig*. Berlin: Akademie Verlag.
- BECKER, HORST (1933): *Mundart und Geschichte im Osterzgebirge*. Halle: Niemeyer. (= Mitteldeutsche Studien 4)

- BECKER, HORST und GUNTHER BERGMANN (1969): Sächsische Mundartenkunde. Entstehung, Geschichte und Lautstand der Mundarten des obersächsischen Gebietes. 2. Auflage. Halle: VEB Max Niemeyer.
- BELLMANN, GÜNTER (1962): Mundart und Umgangssprache in der Oberlausitz. Sprachgeographische Untersuchungen zwischen Schwarzwasser und Lausitzer Neiße. Marburg: Elwert. (= Deutsche Dialektgeographie 62)
- BOERSMAA, PAUL and DAVID WEENINK (2013): Praat: doing phonetics by computer [Computer program]. Version 5.3.41. online. [<http://www.praat.org>, zuletzt abgerufen am 26.09.2013]
- BIRKNER, KARIN und PETER GILLES (2008): "Dialektstilisierung im Reality-Fernsehen". In: CHRISTEN, HELEN und EVELYN ZIEGLER (Hg.): Sprechen, Schreiben, Hören. Zur Produktion und Perzeption von Dialekt und Standardsprache zu Beginn des 21. Jahrhunderts. Wien: Praesens: 101–130.
- DI PAOLO, MARIANNA and MALCAH YAEGER-DROR (eds.) (2011): Sociophonetics. A student's guide. London: Routledge.
- FLEISCHER, WOLFGANG (1961): Namen und Mundart im Raum von Dresden. Toponymie und Dialektologie der Kreise Dresden-Altstadt und Freital als Beitrag zur Sprach- und Siedlungsgeschichte. Berlin: Akademie-Verlag. (= Deutsch-slawische Forschungen zur Namenkunde und Siedlungsgeschichte 11)
- FRINGS, THEODOR (1932): Sprache und Siedlung im mitteldeutschen Osten. Leipzig: Hirzel. (= Berichte über die Verhandlungen der Sächsischen Akademie der Wissenschaften zu Leipzig, Philologisch-Historische Klasse, Bd. 84,6)
- FRINGS, THEODOR (1936): Die Grundlagen des Meißenischen Deutsch. Ein Beitrag zur Entstehungsgeschichte der deutschen Hochsprache. Halle: Niemeyer.

- FRINGS, THEODOR (1956): *Sprache und Geschichte*. Halle: Niemeyer.
- GACKSTATTER, MARJA und OLIVER NIEBUHR (2012): "Eine kontrastive phonetische Analyse niederdeutscher Langvokale". *Linguistik online* 53: 23–53.
- GENDROT, CÉDRIC and MARTINE ADDA-DECKER (2005): "Impact of duration on F1 / F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German". In: *Interspeech 2005*. Lisboa: 2453–2456.
- GERICKE, INGEBORG (1963): "Die Intonation der Leipziger Umgangssprache". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 16: 337–369.
- GILLES, PETER (2005): *Regionale Prosodie im Deutschen. Variabilität der Intonation von Abschluss und Weiterweisung*. Berlin, New York: de Gruyter. (= *Linguistik – Impulse & Tendenzen* 6)
- GÖSCHEL, JOACHIM (1974): *Strukturelle und instrumental-phonetische Untersuchungen zur gesprochenen Sprache. Dargestellt an mitteldeutschen und niederdeutschen Dialekten*. Berlin / New York: de Gruyter. (= *Studia Linguistica Germanica* 9)
- GRICE, MARTINE und STEFAN BAUMANN (2002): "Deutsche Intonation und GToBI". *Linguistische Berichte* 191: 267–298.
- GROSSE, RUDOLF (1955): *Die meißnische Sprachlandschaft: dialektgeographische Untersuchungen zur Obersächsischen Sprach- und Siedlungsgeschichte*. Halle (Saale): Niemeyer. (= *Mitteldeutsche Studien* 15)
- HARRINGTON, JONATHAN (2010a): "Acoustic phonetics". In: HARDCASTLE, WILLIAM, JOHN LAVER und FIONA E. GIBBON (eds.): *The Handbook of Phonetic Sciences*. 2. ed. Chichester: Wiley-Blackwell: 81–129. (= *Blackwell Handbooks in Linguistics*)

- HARRINGTON, JONATHAN (2010b): *The Phonetic Analysis of Speech Corpora*. Chichester: Wiley-Blackwell.
- HEIKE, GEORG und FRANK SCHINDLER (1970): "Versuche zur Dialektgeographie akustisch-phonetischer Messwerte am Beispiel schlesischer Mundartaufnahmen". *Zeitschrift für Dialektologie und Linguistik* 37: 26–43.
- HERRGEN, JOACHIM und JÜRGEN ERICH SCHMIDT (1985): "Systemkontrast und Hörerurteil. Zwei Dialektalitätsbegriffe und ihre entsprechenden Meßverfahren". *Zeitschrift für Dialektologie und Linguistik* 52: 20–42.
- HIRSCHFELD, URSULA (1999): "Phonetische Merkmale des Sächsischen und das Fach Deutsch als Fremdsprache". In: Skibitzki, Bernd und Barbara Wotjak (Hg.): *Linguistik und Deutsch als Fremdsprache – Festschrift für Gerhard Helbig zum 70. Geburtstag*. Tübingen: Max Niemeyer: 109–120.
- HÜNECKE, RAINER, EVELYN KOCH und XAVER KOCH (2012): "Wie hört sich das gegenwärtige Sächsische im Fernsehen an?" In: HÜNECKE, RAINER und KARLHEINZ JAKOB (Hg.): *Die obersächsische Sprachlandschaft in Geschichte und Gegenwart*. Heidelberg: Winter: 143–173.
- IIVONEN, ANTTI (1987): "Zur regionalen Variation der betonten Vokale im gehobenen Deutsch. Kontrastive Evidenz unter besonderer Berücksichtigung des Ostmitteldeutschen und Wienerdeutschen". In: Kahlas-Tarkka, Leena (Hg.): *Neophilologica Fennica*. Helsinki: 87–119. (= *Mémoires de la Société Néophilologique de Helsinki XLV*)
- IIVONEN, ANTTI (1989): "Regional differences in the realization of Standard German vowels". *Helsingin yliopiston fonetiikan laitoksen monisteita* 15: 4–20.
- IIVONEN, ANTTI (1994): "Zur gehobenen regionalen phonetischen Realisierung des Deutschen". In: Viereck, Wolfgang (Hg.): *Verhandlungen des*

- Internationalen Dialektologenkongresses Bamberg 1990, Bd. 3. Stuttgart: Steiner: 311–330. (= Zeitschrift für Dialektologie und Linguistik Beihefte 76)
- JACEWICZ, EWA, ROBERT ALLEN FOX and JOSEPH SALMONS (2007): "Vowel space areas across dialects and gender". In: Proceedings of the ICPHS XVI. Saarbrücken: 1465–1468.
- KEHREIN, ROLAND (2012): Regionalsprachliche Spektren im Raum. Zur linguistischen Struktur der Vertikale. Stuttgart: Franz Steiner. (= Zeitschrift für Dialektologie und Linguistik Beihefte 152)
- KHAN, SAMER UD DOWLA and CONSTANZE WEISE (2013): "Upper Saxon (Chemnitz dialect)". Journal of the International Phonetic Association 43: 231–241.
- KRECH, EVA-MARIA, EBERHARD STOCK, URSULA HIRSCHFELD und LUTZ CHRISTIAN ANDERS (2009): Deutsches Aussprachewörterbuch. Berlin, New York: de Gruyter.
- KÜGLER, FRANK (2007): The intonational phonology of Swabian and Upper Saxon. Tübingen: Max Niemeyer. (= Linguistische Arbeiten 515)
- LABOV, WILLIAM, SHARON ASH and CHARLES BOBERG (2006): Atlas of North American English: phonetics, phonology and sound change: a multimedia reference tool. Berlin: Mouton de Gruyter.
- LAMELI, ALFRED (2004): Standard und Substandard. Regionalismen im diachronen Längsschnitt. Stuttgart: Steiner. (= Zeitschrift für Dialektologie und Linguistik. Beiheft 128)
- LAMELI, ALFRED (2013): Strukturen im Sprachraum. Analysen zur arealtypologischen Komplexität der Dialekte in Deutschland. Berlin / Boston: De Gruyter. (= Linguistik – Impulse & Tendenzen 54)

- LENZ, ALEXANDRA N. (2003): Struktur und Dynamik des Substandards. Eine Studie zum Westmitteldeutschen (Wittlich / Eifel). Stuttgart: Franz Steiner. (= ZDL-Beiheft 125)
- LERCHNER, GOTTHARD (2003): "Aspekte einer Sprachgeschichte des Ostmitteldeutschen". In: BESCH, WERNER, ANNE BETTEN, OSKAR REICHMANN und STEFAN SONDEREGGER (Hg.): Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung. 2. Auflage. Berlin, New York: Walter de Gruyter: 2744–2767.
- LOBANOV, BORIS M. (1971): "Classification of Russian Vowels Spoken by Different Speakes". *Journal of the Acoustical Society of America* 49: 606–608.
- MADDIESON, IAN (1997): "Phonetic Universals". In: HARDCASTLE, WILLIAM J. and JOHN LAVER (eds.): *The Handbook of Phonetic Sciences*. Cambridge: Blackwell: 619–639. (= Blackwell Handbooks in Linguistics 5)
- MOOSHAMMER, CHRISTINE and CHRISTIAN GENG (2008): "Acoustic and articulatory manifestations of vowel reduction in German". *Journal of the International Phonetic Association* 38 / 2: 117–136.
- MOOSMÜLLER, SYLVIA (2007): *Vowels in Standard Austrian German. An Acoustic-Phonetic and Phonological Analysis*. Wien. (= Habilitationsschrift Wien)
- [http://www.kfs.oeaw.ac.at/publications/habil_2007may28_tableofcontents_zus_final.pdf, zuletzt abgefragt am 26.09.2013]
- NICOLAIDIS, KATERINA (2003): "Acoustic Variability of Vowels in Greek Spontaneous Speech". In: *Proceedings of the ICPhS, Barcelona, Spain. August 2-9, 2003*. Barcelona: 3221–3224.
- [http://www.enl.auth.gr/phonlab/Nicolaidis_spon_ac.pdf, zuletzt abgefragt am 26.09.2013]

- PROTZE, HELMUT (1957): Das Westlausitzische und Ostmeißnische: dialektgeographische Untersuchungen zur lausitzisch-obersächsischen Sprach- und Siedlungsgeschichte. Halle (Saale): Niemeyer. (= Mitteldeutsche Studien 20)
- PROTZE, HELMUT (1997): Wortatlas der städtischen Umgangssprache. Zur territorialen Differenzierung der Sprache in Mecklenburg-Vorpommern, Brandenburg, Berlin, Sachsen-Anhalt, Sachsen und Thüringen. Köln u. a.: Böhlau.
- PURSCHE, CHRISTOPH (2012): Regionalsprache und Hörerurteil. Grundzüge einer perzeptiven Variationslinguistik. Stuttgart: Steiner. (= Zeitschrift für Dialektologie und Linguistik Beiheft 149)
- RAMERS, KARL HEINZ (1988): Vokalquantität und -qualität im Deutschen. Tübingen: Niemeyer.
- RUES, BEATE, BEATE REDECKER, EVELYN KOCH et al. (2007): Phonetische Transkription des Deutschen. Ein Arbeitsbuch. Tübingen: Narr. (= Narr Studienbücher)
- SAS INSTITUTE INC. (2009): JMP[®] 8 Statistics and Graphics Guide. Second Edition. Cary, NC: SAS Institute Inc.
- SCHAUFUß, ANJA (i. Dr.): Regionalsprachlichkeit von Sprechern des Obersächsischen im Dreieck Dresden, Chemnitz, Leipzig. Hörerurteil und phonetischer Abstand im diatopischen Vergleich. In: Tagungsband zum Kongress der Internationalen Gesellschaft für Dialektologie des Deutschen e.V. in Kiel September 2012.
- SCHAUFUß, ANJA und BEAT SIEBENHAAR (2012): Spracheinstellungen und phonetische Variation als Ausdruck verschwommener Dialektabgrenzung. In: Zeitschrift für Literaturwissenschaft und Linguistik LiLi 166: 88–109.

- SCHMIDT, JÜRGEN E. (2012): "Warum ist das Obersächsische ein ostmitteldeutscher Dialekt? Die deutschen Sprachlandschaften in den Dokumentationswerken der Dialektgeographie des 19. bis 21. Jahrhunderts". In: HÜNECKE, RAINER und KARLHEINZ JAKOB (Hg.): Die obersächsische Sprachlandschaft in Geschichte und Gegenwart. Heidelberg: Winter.
- SCHMIDT, JÜRGEN ERICH und JOACHIM HERRGEN (2011): Sprachdynamik. Eine Einführung in die moderne Regionalsprachenforschung. Berlin: Erich Schmidt. (= Grundlagen der Germanistik 49)
- SELTING, MARGRET (2001): Dresdener Intonation: Fallbögen. Konstanz. (= InLiSt - Interaction and Linguistic Structures 29)
[<http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-9104>, zuletzt abgerufen am 26.09.2013]
- SELTING, MARGRET (2003): "Treppenkonturen im Dresdenerischen". Zeitschrift für germanistische Linguistik 31: 1–43.
- SIMPSON, ADRIAN P. (1998): Phonetische Datenbanken des Deutschen in der empirischen Sprachforschung und der phonologischen Theoriebildung. Kiel. (= AIPUK 33)
- SENDLMEIER, WALTER F. und JULIA SEEBODE (o. J. [2006]): Formantkarten des deutschen Vokalsystems. online. (www.kw.tu-berlin.de/fileadmin/a01311100/Formantkarten_des_deutschen_Vokalsystems_01.pdf, zuletzt abgerufen am 11.02.2014)
- SPIEKERMANN, HELMUT (2000): Silbenschnitt in deutschen Dialekten. Tübingen: Niemeyer. (= Linguistische Arbeiten 425)
- THOMAS, ERIK R. (2011): Sociophonetics. An Introduction. Basingstoke: Palgrave.

- THOMAS, ERIK R. and TYLER KENDALL (2007–2010): NORM: The vowel normalization and plotting suite. [Online Resource: <http://ncslaap.lib.ncsu.edu/tools/norm/>, zuletzt abgerufen am 26.09.2013]
- VAN DER HARST, SANDER (2011): *The Vowel Space Paradox. A Sociophonetic Study on Dutch*. Utrecht: LOT.
- WALLRAFF, UTA (2007): *Ausgewählte phonetische Analysen zur Umgangssprache der Stadt Halle an der Saale*. Halle. (= Dissertation der Martin-Luther-Universität Halle-Wittenberg)
- WIESINGER, PETER (1983a): "Die Einteilung der deutschen Dialekte". In: BESCH, WERNER, ULRICH KNOOP, WOLFGANG PUTSCHKE und HERBERT ERNST WIEGAND (Hg.): *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung*. Zweiter Halbband. Berlin / New York: de Gruyter: 807–900.
- WIESINGER, PETER (1983b): "Rundung und Entrundung, Palatalisierung und Entpalatalisierung, Velarisierung und Entvelarisierung in den deutschen Dialekten". In: BESCH, WERNER, ULRICH KNOOP, WOLFGANG PUTSCHKE und HERBERT ERNST WIEGAND (Hg.): *Dialektologie. Ein Handbuch zur deutschen und allgemeinen Dialektforschung*. Zweiter Halbband. Berlin, New York: de Gruyter: 1101–1105.
- WREDE, BRITTA, GERNOT A. FINK and GERHARD SAGERER (2000): "Influence of Duration on Static and Dynamic Properties of German Vowels in Spontaneous Speech". In: *International Conference on Spoken Language Processing*, vol 1. Beijing: 82–85.
- Wörterbuch der obersächsischen Mundarten (4 Bände). (1994–2003): Berlin: Akademie Verlag.

Zusammenfassung

Nach dem Abbau der traditionellen Dialekte in zentralen Teilen Sachsens fehlen immer noch gesicherte Resultate zur aktuellen Ausgestaltung der Variationsspektrums. Die punktuellen Untersuchungen von HÜNECKE et al. (2012), KEHREIN (2012), SCHAUFUß / SIEBENHAAR (2012) ergänzt der Aufsatz durch eine weitere punktuelle, diesmal instrumentalphonetische Untersuchung der Stadt Leipzig. Für die Darstellung der Vokalvariation in Leipzig werden Formanten und Länge der Monophthonge von zwölf Frauen untersucht. Die unterschiedlichen Aufnahmesituationen – Interview, Lesetext und intendierter Dialekt – zeigten deutlich unterschiedlich strukturierte Artikulationsräume. Im Vergleich zu anderen Aufnahmen zeigt sich auch in der intendierten Standardsprache eine relativ starke Zentralisierung, besonders bei den Langvokalen. Die Änderung der Lautqualität im intendierten Dialekt und in der Spontansprache gegenüber der Leseaussprache zeigen je Tendenzen zur Zentralisierung, allerdings werden je andere Vokalgruppen zentralisiert: Während in der Spontansprache v. a. die Hinterzungenvokale nach vorne verlagert, also zentralisiert werden, ist im intendierten Dialekt besonders eine Rückverlagerung, d. h. eine Zentralisierung der Vorderzungenvokale zu finden, wobei [i:] stabil bleibt. Dieses Backing der Vorderzungenvokale und der A-Laute sowie die Senkung der kurzen Hochzungenvokale sind Merkmale, die eingesetzt werden, um Mundart zu artikulieren. Da die Vokale im intendierten Dialekt sogar länger sind als in der Lesesprache, kann die Zentralisierung nicht mit Undershoot erklärt werden, sondern deutet auf eine bewusst andere Gestaltung hin. Die Veränderungen der Spontansprache sind dagegen eher als Zentralisierungen, die sich auf das System der Lesesprache beziehen, zu verstehen. Dabei tendiert das stark zentralisierte System der Kurzvokale zu

einer Verteilung der Laute, welche den reduzierten Raum optimal ausnutzt. Neben der Untersuchung der Gestalt der Vokalräume wurde auch deren Größe berechnet und die Veränderung zwischen den Erhebungssituationen verglichen. Deutlich ist die Verkleinerung zwischen Leseaussprache und den beiden andern Situationen, die sich weniger voneinander unterscheiden, allerdings ist die individuelle Variation sehr groß. Auffällig ist aber, dass der Artikulationsraum der Kurzvokale in der Spontansprache einheitlich klein ist und auch bei langsameren Sprecherinnen nicht vergrößert erscheint, so dass hier eine Grenze der Reduktion für diese Situation erreicht scheint. KEHREINS Untersuchung (2012), die in Dresden eine kaum vorhandene Differenzierung der Sprechlagen dokumentiert, kann an dieser Stelle für Leipzig teilweise bestätigt werden. Allerdings ist nicht eindeutig zu entscheiden, ob die Unterschiede zwischen Dresden und Leipzig als regional bedingt zu werten sind oder ob dies auf die andere Erhebungsmethode – ohrenphonetische Transkription und Dialektalitätsmessung nach HERRGEN / SCHMIDT (1985) vs. instrumentalphonetische Messung – zurückzuführen ist. Der instrumentalphonetische Zugang erweist sich aber grundsätzlich als effektive Methode um die variationslinguistisch-dialektologischen Ergebnisse zu überprüfen und zu differenzieren.

Beat Siebenhaar
Universität Leipzig
Institut für Germanistik
Beethovenstraße 15
04107 Leipzig
siebenhaar@uni-leipzig.de