

Zur Vokalquantität in der Mundart der Stadt Zürich

Stephan Schmid (Zürich)

Abstract

This paper analyses the vowel system of the dialect spoken in the city of Zürich, which shows a systematic use of distinctive length contrasts. In order to ascertain the acoustic correlates of the quantity feature, an experimental study has been done, during which three speakers produced a corpus of carefully designed read speech. The duration of 162 vowels and 162 consonants has been measured, as well as the first two formants of 54 high vowels.

The results indicate duration as the only relevant correlate of vowel quantity in this dialect, with a mean V/V: ratio of 0.56. The duration of the following consonants does not vary significantly, nor are there any striking spectral differences between short and long high vowels. From a look at some intensity contours it appears that short vowels are not 'cut' - a finding in line with a recent areal typology of German dialects.

1 Einleitung

Gegenstand dieses Beitrags ist die Quantität - d.h. die phonologisch relevante Dauer - betonter Vokale im Zürichdeutschen. Insbesondere wird dabei untersucht, in welchem durchschnittlichen Zeitverhältnis Kurz- und Langvokale zueinander stehen und inwieweit somit die Länge vokalischer Segmente als phonetisches Korrelat einer phonologischen Quantitätsopposition gelten kann. Daneben werden aber noch zwei weitere, damit verbundene Fragen angesprochen - nämlich, ob bei kurzen und langen Hochvokalen auch ein Unterschied in der Klangfarbe besteht und ob sich die Quantitätsopposition allenfalls auch im Intensitätsverlauf der Vokale (d.h. durch zwei verschiedene Arten des 'Silbenschnitts') manifestiert.

Der Aufsatz ist folgendermassen gegliedert: nach einigen Vorbemerkungen im Abschnitt 2 über die Begriffe der Quantität, der Länge und der Dauer wird im Abschnitt 3 das Vokalsystem des Zürcher Dialekts skizziert und kurz erörtert. Abschnitt 4 ist der Präsentation der Untersuchung sowie der Diskussion einiger methodologischer Probleme gewidmet, während im Abschnitt 5 schliesslich die Resultate instrumentalphonetischer Messungen der Dauer von Vokalen und Konsonanten sowie von Vokalformanten und Intensitätsverläufen

vorgestellt werden. Die Schlussbemerkungen im Abschnitt 6 fassen die Ergebnisse zusammen.¹

2 Definition der linguistischen Begriffe 'Quantität', 'Länge' und 'Dauer'

Die Dauer eines konkret geäußerten Vokals wird durch eine Reihe von Faktoren bestimmt, zum Beispiel durch das individuelle Sprechtempo oder durch performanzbedingte Längungsphänomene (Zögern, Emphase usw.). Daneben gibt es aber auch systematische Dauerunterschiede, die in der einen oder anderen Form in den meisten (wenn nicht in allen) Sprachen und Dialekten vorkommen und zum Teil durch universelle phonetische Prozesse bedingt sind. Nicht in allen Sprachen jedoch hat der Unterschied zwischen kurzen und langen Vokalen einen phonologischen (distinktiven) Wert.

Die folgenden drei Abschnitte sollen einen kurzen Überblick über phonetische und phonologische Aspekte der Vokaldauer geben und dabei insbesondere die drei Begriffe 'Quantität', 'Länge' und 'Dauer' klären.

2.1 Quantität

Mit dem Begriff 'Quantität' ist die relative Dauer von Sprachlauten gemeint, sofern diese phonemisch bzw. phonologisch distinktiv ist. Damit von Quantität gesprochen werden kann, muss eine Sprache also Minimalpaare aufweisen, deren bedeutungsunterscheidende Funktion durch die unterschiedliche Dauer von Segmenten mit gleicher Klangfarbe signalisiert wird. Quantität ist meist eine binäre Kategorie, d.h. man unterscheidet in der Regel zwischen 'langen' und 'kurzen' Vokalen oder Konsonanten; nur für wenige Sprachen wird eine dreifache Quantitätskorrelation angenommen (vgl. Lehiste 1970: 45-48, Fox 2000: 42-46). Als Beispiel für eine Sprache mit Vokalquantität mag das bünderromanische Vallader von Scuol in (1) dienen (Gredig 2000: 49), während die Konsonantenquantität in (2) durch das Ungarische illustriert wird (Fox 2000: 20):

- (1) Vallader /lat/ "Milch" ~ /lat:/ "breit"
 (2) Ungarisch /hal/ "Fisch" ~ /hal:/ "(er/sie) hört"

Unter typologischen Gesichtspunkten muss die Abwesenheit von Quantitätsunterschieden als der unmarkierte, die distinktive Länge von Sprachlauten hingegen als der markierte Fall betrachtet werden. Diese Feststellung trifft insbesondere für die konsonantische Quantität zu, während die Vokalquantität in einer ansehnlichen Minderheit von Sprachen vorhanden ist. So registriert das 317 Sprachen umfassende *UCLA Phonological Segment Inventory Database*

¹ Die der Untersuchung zugrundeliegenden Daten wurden im Rahmen eines Phonetischen Arbeitskreises zum Thema 'Prosodie' im Wintersemester 2001/2002 an der Universität Zürich erhoben. Mein Dank gebührt den drei Sprechern, welche im Abschnitt 4.1 kurz vorgestellt werden, sowie den Teilnehmerinnen und Teilnehmern an dieser Lehrveranstaltung. Die Ergebnisse der Arbeit wurden anlässlich der '2. Tage der Schweizer Linguistik' am 9. November 2002 an der Universität Bern vorgestellt; Astrid Kraehenmann, Hans-Peter Schifferle und Beat Siebenhaar haben bei dieser Gelegenheit interessante Beiträge zur Diskussion geliefert. Wertvolle Verbesserungsvorschläge verdanke ich schliesslich zwei anonymen Gutachtern; für alle Unzulänglichkeiten dieses Beitrags bin ich aber alleine verantwortlich.

(UPSID) die Vokalquantität in 62 Sprachen, d.h. in 19.6% (Maddieson 1984: 129), wobei die Wahrscheinlichkeit für das Auftreten von Quantitätsunterschieden bei einer grösseren Anzahl von Phonemen im Vokalinventar zunimmt (vgl. auch Schmid 1999b: 260-262); die distinktive Konsonantenlänge scheint hingegen weniger verbreitet zu sein (Laver 1994: 436). Funktional lässt sich die grössere Häufigkeit der Quantität bei Vokalen mit der höheren perzeptiven Salienz (im Sinne von Schallfülle und Klangfarbe) des Silbenkerns erklären. Nur wenige Sprachen weisen schliesslich sowohl Vokal- als auch Konsonantenquantität auf - darunter das Finnische, das Estnische, das Maltesische und das Latein (Laver 1994: 443-445).

In einem germanistischen Kontext muss auf die mögliche Korrelation von Vokallänge und Klangfarbe hingewiesen werden. So ist in Sprachen wie dem Englischen die Dauer des Vokals eine redundante Begleiterscheinung der Gespanntheitskorrelation, die artikulatorisch auf einem unterschiedlichen Öffnungsgrad beruht, wie die Minimalpaare (3) und (4) zeigen:

- (3) /fi:l/ "fühlen" ~ /fil/ "füllen"
 (4) /pu:l/ "(Schwimm-)becken" ~ /pul/ "ziehen"

Ähnlich funktioniert im Prinzip des Standarddeutsche, zumindest im Bereich der Hochvokale:

- (5) /'mi:tə/ *Miete* ~ /'mitə/ *Mitte*
 (6) /'ʃpu:kŋ/ *spuken* ~ /'ʃpukŋ/ *spucken*

Allerdings geht im Deutschen bei zwei Vokalen, nämlich bei [ɛ] und [a], die Vokaldauer nicht mit einem Unterschied in der Klangfarbe einher:²

- (7) /'bɛ:tə/ *bäte* ~ /'betə/ *bette*
 (8) /la:m/ *lahm* ~ /lam/ *Lamm*

2.2 Länge

Mit dem Begriff 'Länge' ist hier eine perzeptive Eigenschaft bestimmter Sprachlaute gemeint. Während also 'Quantität' eine phonologische Kategorie darstellt, gehört der Begriff der 'Länge' eher in die auditive Phonetik oder in die Wahrnehmungspsychologie.

Phonologisch gesehen gibt es auch eine Länge, die nicht distinktiv, sondern regelgeleitet ist. Eine Sprache mit allophonischer Vokallänge ist z.B. das Italienische: hier werden alle Vokale gelängt, sofern sie in einer nicht-finalen, betonten und offenen Silbe vorkommen (Schmid 1999a: 165):

- (9) [+vokalisch] → [+lang] / __ \$ (wobei \$ ≠ #)
 /pa.ne/ → [pa:ne] "Brot"

² Vgl. Kohler (²1995: 169-170); für eine ausführlichere Behandlung des Vokalsystems der deutschen Standardsprache siehe Ramers (1988) und Becker (1998).

2.3 Dauer

Die 'Dauer' ist schliesslich das akustische Korrelat von Quantität und/oder Länge, welches im Sprachsignal - meist in Millisekunden - gemessen wird. Die absolute Dauer von Vokalen variiert in der Regel zwischen 30 und 300 ms, wobei Unterschiede von 10-40 ms wahrnehmbar sind (Lehiste 1970: 13). Die effektive Dauer von Sprachlauten hängt dabei von einer ganzen Reihe von Faktoren ab, von denen die durch die kommunikative Absicht der Sprecher bedingte phonologische Quantität nur einen darstellt. Nebst offensichtlichen Performanzerscheinungen wie z.B. der individuellen Sprechgeschwindigkeit gibt es auch Tendenzen, die offensichtlich durch die Natur der Sprachproduktion bedingt sind, wobei zwischen 'intrinsischer' und 'extrinsischer' Dauer unterschieden werden muss.

Das Phänomen der 'intrinsischen' Vokaldauer ist in verschiedenen Sprachen beobachtet worden und stellt vermutlich ein phonetisches Universale dar. Gemäss einer allgemeinen Tendenz sind Tiefvokale länger als Hochvokale, was mit der grösseren Bewegung der Artikulatoren im Übergang von Konsonanten zu offenen Vokalen zu erklären ist (Lehiste 1970: 18-19).

Die 'extrinsische' Dauer von vokalischen Segmenten hängt z.B. von den phonetischen Eigenschaften der vorhergehenden oder nachfolgenden Konsonanten ab, wobei sowohl der Artikulationsort als auch die Artikulationsart eine Rolle spielen (Lehiste 1970: 19-30). So sind Vokale vor bilabialen Konsonanten oft kürzer als vor Konsonanten mit anderer Artikulationsstelle, da offenbar die kleineren Verschlussbewegungen der Lippen schneller erfolgen als die grösseren und schwerfälligeren Bewegungen des Zungenkranzes oder des Zungenrückens. Allfällige Abhängigkeiten der Vokaldauer von der Artikulationsart des folgenden Konsonanten scheinen sprachspezifisch zu sein; eine universale Tendenz kann allenfalls bezüglich der Stimmhaftigkeit festgestellt werden, indem Vokale vor stimmhaften Konsonanten länger sind als vor stimmlosen.

Die extrinsische Vokaldauer wird aber auch durch Faktoren höherer Ordnung bestimmt, insbesondere durch die phonotaktische Struktur der Silbe, die Anzahl der Silben im Wort und die Position des Worts im prosodischen Verlauf der Äusserung. In offener Silbe sind Vokale normalerweise länger als in geschlossener Silbe (Laver 1994: 447); die allophonische Regel (9) des Italienischen stellt somit nur die sprachspezifische Phonologisierung einer universellen phonetischen Tendenz dar. Ebenso hat man in verschiedenen Sprachen festgestellt, dass die durchschnittliche Dauer der Vokale in dem Masse abnimmt wie die Anzahl der Silben in einem Wort zunimmt: Vokale in Einsilblern sind länger als Vokale in Zweisilblern, die wiederum länger sind als Vokale in Dreisilblern usw. (Lehiste 1970: 40). Schliesslich zeigen verschiedene Sprachen das Phänomen der präpausalen Dehnung, wodurch das Sprechtempo gegen Ende einer Äusserung verlangsamt und die Dauer insbesondere der letzten Silbe verlängert wird (Laver 1994: 532-533).

Was nun die effektive Umsetzung der phonologischen Quantität im Sprachsignal anbelangt, so scheint es gewisse Unterschiede zwischen den einzelnen Sprachen zu geben. Als Mass wird häufig die prozentuale Dauer der Kurzvokale im Verhältnis zur Dauer der Langvokale angegeben. Für das Finnische beträgt dieses Verhältnis z.B. 0.44, für das Dänische 0.50 und

für das Serbokroatische 0.67 (Lehiste 1970: 34). Mit anderen Worten: im Finnischen ist der Unterschied zwischen Kurz- und Langvokalen am ausgeprägtesten, im Dänischen dauert ein Kurzvokal im Durchschnitt halb so lang wie ein Langvokal, während der Dauerunterschied im Serbokroatischen am wenigsten deutlich ist. Bei solchen Angaben handelt es sich jeweils um Mittelwerte von Dauerverhältnissen, die je nach Sprecher unterschiedlich ausfallen können: bei den vier von Ramers (1988: 197-200) untersuchten Sprechern des Deutschen beträgt z.B. der prozentuale Anteil der Kurzvokale an der Dauer der Langvokale - unabhängig von ihrer Klangfarbe (vgl. 2.1) - 0.39, 0.48, 0.50 und 0.61.

3 Das Vokalsystem des Zürichdeutschen

In der vorliegenden Untersuchung wird für die Mundart der Stadt Zürich folgendes Vokalsystem angenommen:

| | | |
|------|------|------|
| i i: | y y: | u u: |
| e e: | ø ø: | o o: |
| ɛ ɛ: | œ œ: | |
| æ æ: | | ɒ ɒ: |

Dieses Inventar von Vokalphonemen stützt sich in erster Linie auf die Introspektion des Schreibenden, der selbst in der Stadt Zürich lebt; es deckt sich aber mit dem System, welches Keller (1961: 36) für den Dialekt der Stadt Winterthur postuliert.

3.1 Vokalqualität

Das Vokalsystem des Stadtzürcher Dialekts enthält vier Öffnungsgrade (hoch, mittelhoch, mitteltief und tief) und drei Artikulationsreihen (vorne ungerundet, vorne gerundet, hinten); aus Gründen der Symmetrie wird [æ], welches im IPA-Trapez als Halbtiefvokal definiert ist, mit dem Tiefvokal [ɒ] in einer Serie zusammengefasst. Die in diesem System angelegten Asymmetrien entsprechen durchaus den aus der phonologischen Typologie bekannten universellen Tendenzen, die ihrerseits in den physiologischen Bedingungen des Vokaltrakts begründet sind; dazu gehören a) die grössere Anzahl der ungerundeten Vordervokale gegenüber den anderen beiden Artikulationsreihen und b) die progressive Abnahme der gerundeten Vordervokale von oben nach unten (Martinet 1962 [1969]: 98, Maddieson 1984: 14).

Dennoch bestehen Fragen zur Struktur dieses Vokalsystems, insbesondere bezüglich der Definition der Klangfarbe einzelner Phoneme. So erachtet Keller (1961: 35, 37) die kurzen Hochvokale /i u y ø/ als eine Besonderheit der Winterthurer Mundart, an deren Stelle im übrigen Kanton Zürich kurze Halbhochvokale vorkämen, wonach ähnliche Verhältnisse gelten würden wie im Deutschen (vgl. 2.1). Im Widerspruch zu dieser Analyse postuliert der gleiche Autor für die Stadt Zürich sogar eine phonologische Opposition /y:/ ~ /Y:/, woraus ein umfangreicheres Vokalinventar resultieren würde; allerdings wird nicht angegeben, wie dieses System genau aufgebaut sein soll. Auch die populärwissenschaftliche Darstellung von Schobinger (2001: 118) geht von fünf Öffnungsgraden aus und liefert einige wenige

Minimalpaare für lange Hoch- und Halbhochvokale, wobei immerhin folgendes vermerkt wird: "D unterschädig vo offne und gschlossne wokaal schiint am undergaa z sii". Ob nun Stadtzürcher für die beiden Bedeutungen "Zeug" und "(Eisenbahn-)Züge" ein homonymes Wort /tʰy:k/ verwenden oder zwei verschiedene Lexeme /tʰy:k/ und /tʰy:k/, kann aufgrund der vorliegenden Daten und des primären Erkenntnisinteresses nicht abschliessend geklärt werden. Trotzdem soll im Abschnitt 5.3 auf den Öffnungsgrad von kurzen Hochvokalen näher eingegangen werden.³

Auf der anderen Seite ist die hier angenommene Rundung des hinteren Tiefvokals scheinbar umstritten. Diese wird zwar nicht nur von Keller (1961: 36) für den Winterthurer Dialekt angesetzt, sondern auch Willi (1996) transkribiert die Tiefvokale seiner aus verschiedenen Stadtvierteln und Kantonsteilen stammenden Sprechern konsequent als [ɒ]. Diese Annahme entspricht durchaus der dialektologischen Tradition, wie sie etwa durch den Sprachatlas der deutschen Schweiz (SDS) repräsentiert wird: so bezeugen die Karten für mhd. *a* und *â* aus dem ersten Band (S. 11 und 61) an den beiden Punkten Zürich (37) und Zürich-Schwamendingen (38) ein "stark verdumpftes a", welches in der Aufstellung von Hofer (1997: XII) mit dem IPA-Symbol [ɒ] gleichgesetzt wird.⁴ Nun entspricht diese Rundung aber nicht unbedingt der Wahrnehmung einzelner Stadtzürcher, welche an dieser Stelle lieber ein ungerundetes [ɑ] sähen.⁵ Möglich ist, dass hier eine freie, sprecherabhängige Variation vorliegt; denkbar wäre aber auch, dass es sich um eine soziolinguistische Variable handelt - etwa im Sinne einer stärkeren Rundung bei Sprechern aus der Landschaft und/oder den unteren sozialen Schichten, der eine schwächere oder sogar fehlende Rundung in der urbanen Varietät der Mittel- bzw. Oberschicht gegenüberstehen würde. Um diese Frage zu beantworten, fehlt zur Zeit jedoch die nötige empirische Evidenz, die nur durch Formantmessungen in relativ grossen Korpora geliefert werden kann.⁶

³ Vokalsysteme mit fünf Öffnungsgraden wurden von Moulton (1973) im Toggenburg und im St. Galler Rheinthal geortet (anhand einer Auswertung der Karten des Sprachatlases der Deutschen Schweiz), während Traunmüller (1982) für die niederösterreichische Mundart experimentalphonetische Evidenz für fünf Höhenstufen anführt. Zumindest phonetisch betrachtet weist auch das Standarddeutsche fünf Öffnungsgrade auf, sofern man Kurz- und Langvokale in einem System zusammenfasst (vgl. Becker 1998: 32) - ein eher seltenes und typologisch markiertes Phänomen (vgl. Martinet 1962 [1969]: 98, Ladefoged/Maddieson 1996: 289-290). Auch in den schweizerdeutschen Dialekten scheinen eher vierstufige Systeme vorzuherrschen: vgl. dazu etwa Dieths (1950: 358) Analyse des Vokalsystems der Mundart von Kesswil (TG), die zusätzlich zu unserem Zürcher Inventar noch die Phoneme /œ ɔ ɔ:/ aufweist.

⁴ Angesichts dieser Datenlage ist schwer zu verstehen, weshalb Christen (2002: 219) die "Verdumpfung von lang und kurz /a/" als Abweichung von der "Grundmundart" ..., wie sie etwa im Sprachatlas der deutschen Schweiz (SDS) ausgewiesen ist" bezeichnet.

⁵ Phonologische Argumente gegen eine Reihenbildung der Tiefvokale /æ ɑ/ und für eine durch das Merkmal [±labial] oder [±gerundet] unterschiedene Reihe /æ ɒ/ finden sich bei Haas (1978: 323-324).

⁶ Akustisch gesehen führt die Rundung eines Vokals zu einer 'Verlängerung' seines Resonanzraums und somit zu einem niedrigeren zweiten Formanten (vgl. Neppert ⁴1999: 150). Stichproben bei den beiden männlichen Sprechern dieser Untersuchung (vgl. 4.3) ergeben in der Tat sehr niedrige F2-Messungen in den Testwörtern [mɔ:nə] und [mɒnə], nämlich 952 Hz und 970 Hz, was im Sprachvergleich eher dem Wert eines [ɔ] als demjenigen eines [ɑ] entspricht (vgl. Delattre 1981: 71, 73). Zu den Vokalformanten siehe auch Fussnote 13.

Zu erwähnen ist schliesslich die Dauer des Vokals [œ], der als einziger keine Quantitätskontraste aufweist und gemäss dem 'Winterthurer System' nach Keller (1961: 36) nur als Langvokal erscheint. Introspektive und auditive Evidenz legen die Annahme eines zugrundeliegenden Langvokals nahe (vgl. [tœ:rf] "darf", [ʒœ:f] "Schafe", [vʁœ:ğə] "fragen" usw.), obwohl aufgrund üblicher Markiertheitsverhältnisse eigentlich eher ein phonemischer Kurzvokal zu erwarten wäre, der allenfalls in gewissen Kontexten allophonisch gedehnt wird.⁷ Immerhin treten im typologischen Vergleich gerundete Vordervokale häufiger als andere Vokale nur lang auf - eine Tendenz, die gerade bei /œ/ ziemlich ausgeprägt zu sein scheint (vgl. Maddieson 1984: 129-130).

3.2 Vokalquantität

Eine grundlegende Eigenschaft des Vokalsystems der Stadtzürcher Mundart bleibt jedoch die systematische Nutzung der distinktiven Vokallänge, die durch folgende Minimalpaare belegt werden kann:

- (10) /i/ ~ /i:/ *sibe* "7" ~ *siibe* "sieben (Verb)"
- (11) /e/ ~ /e:/ *Bett* ~ *Beet*
- (12) /ɛ/ ~ /ɛ:/ *Herr* ~ *hèèr* "her"
- (13) /æ/ ~ /æ:/ *gäll* "gell" ~ *gääl* "gelb"
- (14) /ɔ/ ~ /ɔ:/ *Mane* "Männer" ~ *maane* "mahnen"
- (15) /o/ ~ /o:/ *hole* "holen" (Verb) ~ *hoole* "hohler" (Adj.)
- (16) /u/ ~ /u:/ *Bruch* ~ *Bruuch* "Brauch"
- (17) /y/ ~ /y:/ *Füli* "Füllfederhalter" ~ *Füüli* "Faulheit"
- (18) /ø/ ~ /ø:/ *Böge* "Bogen (pl.)" ~ *Böögge* "Narren"

Diese Beispiele zeigen auf eindrückliche Art und Weise, wie konsequent das Merkmal der Quantität das Vokalsystem des Zürichdeutschen prägt. Es ist nicht schwer, weitere Minimalpaare für die einzelnen Klangfarben zu finden, z.B.

- für [i] *Wise* "Wiese" ~ *wiise* "weisen",
- für [y] *Müli* "Mühle" ~ *Müüli* "Mäulchen" und *füre* "hervor" ~ *füüre* "feuern",
- für [u] *Brune* "Brunnen" ~ *bruune* "brauner" und *sure* "surren" ~ *suure* "saurer",
- für [ɔ] *schare* "scharren" ~ *Schaare* "Scharen" und *zale* "bezahlen" ~ *Zaale* "Zahlen" (vgl. auch Hotzenköcherle 1986: 321-322).

⁷ Die hier verwendete phonetische Transkription der zürichdeutschen Konsonanten folgt im wesentlichen der von Eugen Dieth (vgl. Fussnote 10) geprägten Tradition. Insbesondere wird für Obstruenten eine phonologische Opposition zwischen *fortes* und *lenes* angenommen, die in Minimalpaaren vom Typus [lɔtə]"Latte" vs. [lɔdə]"Laden" in Erscheinung tritt (vgl. Dieth 1950: 357-360). Im akustischen Signal manifestiert sich dieser Unterschied hauptsächlich als Unterschied der Konsonantendauer, wie bereits von Dieth (1950: 180) bemerkt und schliesslich von Willi (1996) umfassend nachgewiesen worden ist. Allerdings sollte die Transkription der *lenes* mit dem diakritischen Zeichen für 'Entsonorisierung' nicht dahingehend interpretiert werden, dass es sich in jedem Fall um ursprünglich stimmhafte Konsonanten handelt (vgl. z.B. [ʒri:ɸə] "schreiben"); gemäss der Diethschen Analyse der schweizerdeutschen Phonotaktik kommen Frikative im Wortanlaut zugrundeliegend einfach nur als *lenes* vor (vgl. Dieth 1950: 360).

Einzig (18) stellt nur ein Quasi-Minimalpaar dar, da neben der Länge des Vokals auch ein phonemischer Kontrast beim nachfolgenden Konsonanten vorliegt.

Die grundlegende Bedeutung der Vokalquantität wird in der dialektologischen Literatur zum Schweizerdeutschen beinahe als selbstverständlich vorausgesetzt; so charakterisiert Keller (1961: 37) etwa den Zürcher Vokalismus als "ten-vowels system based on quantity distinction", ohne diese Aussage jedoch durch Minimalpaare nachzuweisen. Die Vokallänge kommt im Übrigen in der Dietschen Dialektschrift (vgl. Abschnitt 4.1) klar zum Ausdruck, da sie auch in anderen schweizerdeutschen Dialekten distinktiv ist.⁸ Wohl aufgrund ihres unbestrittenen phonemischen Status war die Dauer der Vokale aber bisher nie Gegenstand spezifischer instrumentalphonetischer Untersuchungen; immerhin liefert Willi (1996: 104) im Rahmen seiner Erforschung der zürichdeutschen Plosive aufschlussreiche Mittelwerte für kurze und lange Tiefvokale (118 ms bzw. 310 ms).

Ziel der vorliegenden Arbeit ist es nun, diese Forschungslücke zu schliessen. Um Angaben zur Dauer der phonemischen Kurz- und Langvokale des Zürichdeutschen zu machen, wurde deshalb am Phonetischen Laboratorium der Universität Zürich ein kleines Korpus von Sprachdaten erhoben, das im nächsten Abschnitt vorgestellt wird.

4 Die Erhebung der Daten

4.1 Methodologische Grundprobleme

Für die Konstruktion einer geeigneten Datenbank stellt sich ein Dilemma, das aus dem unauflösbaren Widerspruch zwischen den beiden Erfordernissen einer maximalen Kontrolle der strukturellen Variablen einerseits und der maximalen Authentizität der Sprachdaten andererseits entsteht.

Eine in der experimentellen Phonetik beliebte Methode verwendet 'Logatome', d.h. Phantasiewörter, mittels derer künstliche 'Minimalpaare' vom Typus *pade* vs. *paade*, *pide* vs. *piide*, *pode* vs. *poode* usw. gebildet werden können. Dieses Prozedere ermöglicht es, den segmentalen und phonotaktischen Kontext der untersuchten Wörter konstant zu halten und somit zwei mögliche Variablen der extrinsischen Vokallänge auszuschliessen (vgl. Abschnitt 2.3). Logatome wurden z.B. von Willi (1996: 85-114, 207-208) in einer seiner Untersuchungen zum *fortis/lenis*-Kontrast im Zürichdeutschen eingesetzt.

Falls bedeutungstragende Wörter einer Sprache verwendet werden, so sucht man oft nach Minimalpaaren, die auf der phonologischen Eigenschaft der Quantität beruhen. Die entsprechenden Wortformen werden in phonetischen Untersuchungen meist in 'Rahmensätze' vom Typus "Ich habe dreimal X gesagt" (Bosoni 1995: 348) oder "Ich habe nicht X, sondern

⁸ Hotzenköcherle (1986: 322) betont dazu allerdings mit Nachdruck, "dass in der Art, wie die vokalischen Quantitätsoppositionen realisiert werden, im Schweizerdeutschen grosse Unterschiede von Mundart zu Mundart bestehen". So hat die Dehnung von alter Kürze in offener Silbe nicht in allen Dialekten gleich konsequent stattgefunden, wie an den Karten des zweiten Bandes des SDS (1965) ersichtlich ist. - Eine vergleichende Studie zur diachronen Entwicklung der Vokallänge im Thurgauer und Bündner Dialekt legt nun Kraehenmann (2003) vor.

Y gesagt" (Willi 1996: 209-210) gebettet, um den prosodischen Kontext gleich zu halten und insbesondere das Phänomen der präpausalen Dehnung (vgl. Abschnitt 2.3) zu vermeiden.

Diese beiden methodischen Vorkehrungen haben also klare Vorteile in Bezug auf die Kontrolle der extrinsischen Vokallänge, weisen aber auf der anderen Seite gravierende Nachteile auf. Die Quantität als phonologisch distinktives Merkmal erfüllt eine wichtige Funktion für die kommunikative Absicht eines Sprechers und kann somit nicht von der Bedeutung der zu untersuchenden Wörter getrennt werden, weshalb Logatome für unsere Zwecke nicht geeignet sind. Dieser Makel kann auch durch die Verwendung von Minimalpaaren nicht behoben werden, solange diese in Rahmensätze mit metasprachlichem Sinn eingebettet sind. Das Vorgehen bleibt im Gegenteil bis zu einem gewissen Grad zirkulär: man beobachtet, wie sich Kurz- von Langvokalen unterscheiden, nachdem man die Sprecher durch die isolierte oder zumindest metasprachliche Nennung der Wörter implizit auf das Phänomen der Quantität hingewiesen hat.

Eine von obigen Methoden radikal verschiedene Erhebung der Daten würde von einem spontansprachlichen Korpus ausgehen, wie es z.B. anhand von Diskussionen in Radio- und Fernsehsendungen gesammelt werden kann. Die Nachteile dieser Vorgehensweise sind ebenso offensichtlich: die Wahrscheinlichkeit, dass das gesuchte Merkmal (hier: die Quantität für neun verschiedene Vokalfarben) in seiner ganzen Ausprägung in den Daten auftaucht, muss als sehr gering betrachtet werden. Bei Gesprächssituationen kommen zusätzlich noch Schwankungen in Tempo und Rhythmus hinzu, die der prosodischen Steuerung der kommunikativen Interaktion dienen, was die Kontrolle der Variablen zusätzlich erschwert (vgl. Auer/Couper-Kuhlen/Müller 1999).

4.2 Die Testsätze

Angesichts dieses methodologischen Dilemmas erscheint die Konstruktion einer Liste von sinnhaften Testsätzen, die von Versuchspersonen gelesen werden, als valable Alternative. Solche Testsätze werden zwar eigens zur Überprüfung des relevanten Merkmals erfunden und stellen somit keinen organischen Text dar; ausserdem unterscheidet sich gelesene Sprache von gesprochener Spontansprache durch eine Reihe prosodischer Erscheinungen, etwa durch die geringere Häufigkeit und kürzere Dauer der Pausen (siehe Soriano 1997: 94 für Vergleichswerte aus dem Italienischen). Da die untersuchten Wörter aber in einem sinnhaften Kontext erscheinen und die Versuchspersonen nicht auf den Zweck der Untersuchung hingewiesen werden, entsteht trotzdem ein relativ 'natürlicher' Sprachgebrauch, der insbesondere den wesentlichen Vorteil hat, die Zirkularität in der Anlage der Untersuchung zu vermeiden.

Aus diesen Gründen erfolgte die Erhebung der Daten anhand von 18 Testsätzen, welche die Wortformen der 9 Minimalpaare in (10)-(18) enthalten. Somit sind alle vokalischen Klangfarben, für welche eine Quantitätsopposition angenommen wird, im Testmaterial enthalten.

- (19) **Sibe** Schtuck han i kauft.
 (20) Bi de Jurischte tüend's schiint's rächt **siibe** bi de Prüefige.
 (21) Im **Bett** isch's mer immer no am wöölschte.
 (22) I dem **Beet** han i nume Salaat pflanzt.
 (23) Hüt am Morge hät en **Herr** für dich aaglütet.
 (24) Vo deet **hèèr** blaast de Wind.
 (25) **Gäll**, du kännsch mi nööd!
 (26) Er isch ganz **gääl** worde vor Niid.
 (27) Zweek wackeri **Mane** händ ere ghulfe.
 (28) Rächtnige **maane** tuen i nöd gern.
 (29) Ich **hole** no schnäll es Päckli Zigarette.
 (30) Es isch emaal en Maa gsii, dèè hät en **hoole** Zaa ghaa.
 (31) Das hät zum **Bruch** gfüert zwüschet öis bäide.
 (32) Dèè **Bruuch** wird bi öis nüme pflägt.
 (33) Ich schriibe lieber mit **Füli** als mit Bleistift.
 (34) Vor luuter **Füüli** hät er natüürli wider nüüt gmacht.
 (35) Zää **Böge** han i müese n uusfüle.
 (36) So vill **Böögge** gseesch nöd emaal a de Fasnacht.⁹

Die Verschriftung der schweizerdeutschen Mundarten stellt nicht nur den Forscher, sondern auch die Sprecher vor ein weiteres Problem. Da der Dialekt fast ausschliesslich in der mündlichen Kommunikation verwendet wird, stellt das Lesen eines geschriebenen Textes für viele Personen eine ungewohnte Aufgabe dar. Auch gibt es keine allgemein anerkannte orthographische Norm, welche z.B. von der Schule vermittelt würde. Die hier verwendete Verschriftung der Testsätze lehnt sich stark an die Dieth-Schreibung an, die 1938 zum ersten Mal vorgeschlagen wurde¹⁰ und - wie der Herausgeber der 2. Auflage (1986: 15) betont - bewusst sehr "lautnah" konzipiert ist: "Der Kern jeder Dieth-Schreibung ist die einfache und eindeutige Kennzeichnung der Vokallänge". In den Richtlinien wird unter dem Punkt 11 (*Die Länge der Vokale*) festgehalten: "Die Länge eines Vokals soll nur durch seine Doppelsetzung ausgedrückt werden. Die Länge soll überall, also auch in Vor- und Nachsilben, angegeben werden" (1986: 37).

Diese Empfehlung ist in der vorliegenden Untersuchung konsequent befolgt worden, nämlich durch die ausnahmslose Verwendung von einfachen Graphemen bzw. von Digraphen für

⁹ Hier eine mehr oder weniger wörtliche Übersetzung der Beispielsätze: (19) Sieben Stück habe ich gekauft. (20) Bei den Juristen scheinen sie recht zu sieben bei den Prüfungen. (21) Im Bett ist es mir immer noch am wohlsten. (22) In diesem Beet habe ich nur Salat gepflanzt. (23) Heute morgen hat ein Herr für dich angerufen. (24) Von dort her bläst der Wind. (25) Gell, du kennst mich nicht. (26) Er ist ganz gelb geworden vor Neid. (27) Zwei wackere Männer haben ihr geholfen. (28) Rechnungen mahnen tue ich nicht gern. (29) Ich hole noch schnell ein Päcklein Zigaretten. (30) Es war einmal ein Mann, der hatte einen hohlen Zahn. (31) Das hat zum Bruch geführt zwischen uns beiden. (32) Dieser Brauch wird bei uns nicht mehr gepflegt. (33) Ich schreibe lieber mit Füllfeder als mit Bleistift. (34) Vor lauter Faulheit hat er natürlich wieder nichts gemacht. (35) Zehn Bogen musste ich ausfüllen. (36) So viele Narren siehst du nicht einmal an der Fastnacht.

¹⁰ Die *Schwyzertütschi Dialäkttschrift* wurde von der Schriftkommission der Neuen Helvetischen Gesellschaft erarbeitet, welche von Prof. Eugen Dieth (1893-1956) - dem Gründer des Phonetischen Laboratoriums der Universität Zürich - geleitet wurde.

Kurz- und Langvokale, wie man den Minimalpaaren (10)-(18) und vor allem den Testsätzen (19)-(36) entnehmen kann.¹¹ Es ist wahrscheinlich, dass dieses Vorgehen die Versuchspersonen implizit dazu verleitet hat, bei der Lektüre kurze und lange Vokale klar zu unterscheiden. Umgekehrt kann man aber argumentieren, dass gerade die relative Leichtigkeit, mit der die Informanten die geschriebenen Sätze gelesen haben, einen Beweis für die psychologische Realität der Quantitätsopposition liefert - jede andere Verschriftung wäre unnatürlich ausgefallen und hätte die Erhebung der Daten zwangsläufig erschwert.

Was die phonotaktische Struktur der Testwörter anbelangt, so handelt es sich einerseits um zweimal vier Einsilbler vom Typus CVC (/b̥et/, /h̥er/, /g̥æɫ/) oder CCVC (/b̥rux/) sowie um zweimal fünf Zweisilbler vom Typus 'CV.CV (/z̥ib̥ə/, /v̥yɫi/, /b̥øg̥ə/, /'holə/, /'mɔnə/). Aufgrund der Erfahrungen mit anderen Sprachen (vgl. 2.3) ist hier mit einer höheren durchschnittlichen Dauer der Vokale in Einsilblern zu rechnen. Auch der segmentale Kontext der untersuchten Wörter variiert beträchtlich: in vier (bzw. acht) Fällen ist der folgende Konsonant ein stimmloser Obstruent (nämlich /b/, /t/, /g̥ k/, /x/ in /z̥ib̥ə/, /b̥et/, /b̥øg̥ə/ /b̥rux/ sowie in den entsprechenden Wörtern mit Langvokal), während in den anderen fünf (bzw. zehn) Fällen auf den Vokal ein Sonorant und somit ein stimmhafter Konsonant folgt (nämlich /n/, /r/, /l/ in /'mɔnə/, /h̥er/ und in /v̥yɫi/, /g̥æɫ/, /'holə/ (bzw. in den entsprechenden Wörtern mit Langvokal). Angesichts der phonotaktischen Variabilität des Testmaterials ist demnach der Faktor der extrinsischen Vokaldauer (vgl. 2.3) bei der Auswertung der Daten zu berücksichtigen.

Keinen direkten Einfluss auf die Vokaldauer dürfte hingegen die syntaktische Kategorie der untersuchten Wörter ausüben; es handelt sich um

- zehn Nomen (*Mane, Bett, Beet, Bruch, Bruuch, Hèrr, Füli, Füüli, Böge, Böögge*),
- drei Verben (*maane, siibe, hole*),
- je ein Adjektiv mit attributiver und prädikativer Funktion (*hoole Zaa, gääl worde*),
- ein Zahlwort (*sibe*),
- ein Richtungsadverb (*hèèr*),
- eine Interjektion (*gäll*).

Für die Vokaldauer relevant ist hingegen die prosodische Position der Testwörter, wobei der Verzicht auf einen fixen Rahmensatz erneut eine gewisse Variabilität einführt. Die meisten Testwörter kommen aber in einer relativ prominenten Stellung und in der ersten Hälfte der Äusserung vor, so dass kaum mit Abschwächungserscheinungen und schon gar nicht mit präpauzaler Dehnung zu rechnen ist; die einzige Ausnahme bildet hier die Interjektion *gäll*, die im Satz (25) tatsächlich vor einem Komma erscheint.

Die Aufnahme der Daten erfolgte im November 2001 im Phonetischen Laboratorium der Universität Zürich. Um eine optimale Klangqualität zu erreichen, lasen drei Versuchspersonen (vgl. 4.3) das Testmaterial in einer schalldichten Tonkabine. Die 18 Sätze wurden den Informanten dreimal (auf drei Seiten) in unterschiedlicher Reihenfolge vorgelegt, wobei die

¹¹ Auf den Richtlinien der Dieth-Schreibung beruht auch die Verwendung eines Gravis-Akzents zur Kennzeichnung von Mitteltiefvokalen, z.B. in *dèè* [de:] "der".

Wortformen der Minimalpaare nie in aufeinander folgenden Sätzen standen. Die Informanten waren nach der Lektüre nicht in der Lage, das Ziel der Untersuchung zu erkennen. Trotz der ungewohnten Aufgabe darf das erhobene Sprachmaterial als relativ natürlich gewertet werden, da die Sprecher z.T. spontan idiosynkratische Varianten des gedruckten Texts verwendeten, z.B. freie Allomorphe von Verbformen (*ich schriib* anstatt *ich schriibe*).

4.3 Die Informanten

Die drei Versuchspersonen wurden alle in der Stadt Zürich geboren, wo sie - zum Teil mit kleineren Unterbrüchen - seither auch gelebt haben.

- Sprecherin A wurde 1956 geboren und ist in Zürich-Höngg aufgewachsen. Die Mutter stammt aus dem Zürcher Oberland, der Vater ist früh gestorben.
- Sprecher B wurde 1941 geboren und ist im Industriequartier aufgewachsen. Der Vater stammt von Hausen a. Albis, die Mutter aus Stäfa.
- Sprecher C wurde 1960 geboren und ist in Zürich-Wollishofen aufgewachsen. Beide Eltern sind Stadtzürcher.

Somit dürften die drei Informanten als typisch gelten für die Stadtzürcher 'Grundmundart', da ihre Sprachbiographien keinen Einfluss anderer Sprachen oder Mundarten aufweisen; beim in der Familie gesprochenen Dialekt handelt es sich in jedem Fall um eine Varietät des Zürich-deutschen.

4.4 Die Auswertung der Daten

Die auf Tonbandkassetten aufgenommenen Sprachdaten wurden mit einem DSP Sonagraph 5500 der Firma Kay Elemetrics digitalisiert und einer elektroakustischen Analyse unterzogen. Der praktische Vorteil dieser Hard- und Software liegt darin, dass die analoge Eingabe des Sprachsignals in Echtzeit digitalisiert wird, was gegenüber anderen vergleichbaren Analysewerkzeugen eine nicht zu unterschätzende Zeitersparnis mit sich bringt.

Im Sprachsignal wurden folgende Parameter gemessen und mittels beschreibender Statistik dargestellt:

- die Dauer der entsprechenden Kurz- und Langvokale (in Millisekunden),
- die Dauer der nachfolgenden Konsonanten (in Millisekunden),
- die Frequenz der ersten beiden Formanten bei den untersuchten Hochvokalen /i y u/ bzw. /i: y: u:/ (in Hertz).

Um die Dauer der Kurz- und Langvokale sowie die nachfolgenden Konsonanten messen zu können, wurden die entsprechenden Laute manuell segmentiert in Breitbandspektrogrammen.¹² Abb. 1 zeigt die Segmentierung des von der Sprecherin A geäußerten Wortes ['zi:bə].

¹² In dieser Form der Signaldarstellung wird eine komplexe Schallwelle in ihre Bestandteile zerlegt, um die Energieverteilung des Obertonbereichs (auf der Ordinate) in ihrem zeitlichen Verlauf (auf der Abszisse) darzustellen. Im Gegensatz zu einem Schmalbandspekrogramm, dessen Filter eben nur ein schmales

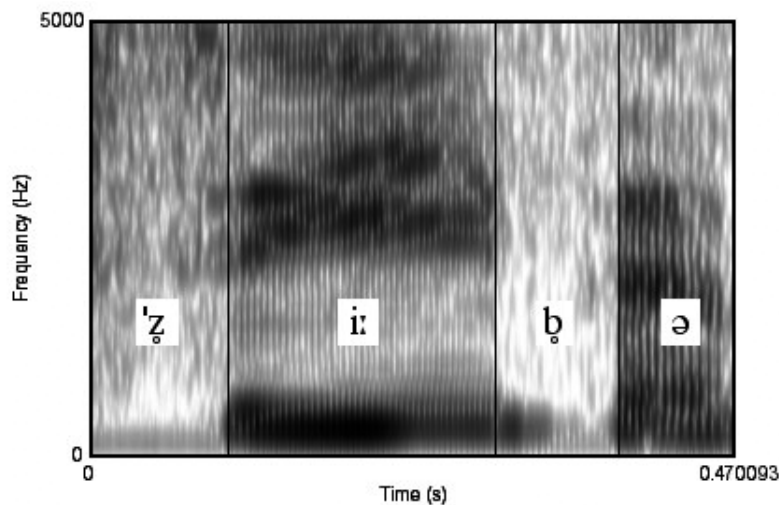


Abb. 1 Segmentierung des Wortes [zi:bə] im Breitbandspektrogramm

Die Segmentierung erfolgte anhand in der phonetischen Literatur verwendeter Standardkriterien (vgl. Schwab et al. 1998). Berücksichtigt wurde insbesondere der Formantenverlauf, also bei Plosiven in erster Linie die Verschlussdauer und nicht eine allfällige Aktivität der Stimmlippen, die in der ersten Phase von stimmlosen Konsonanten noch als Stimmhaftigkeitsbalken im unteren Teil des Spektrogramms erscheinen kann. Bei schwierigeren Segmentierungsproblemen, etwa bei der Abgrenzung von Liquiden und Nasalkonsonanten von den vorhergehenden Vokalen, wurde immer auch eine auditive Überprüfung der segmentierten Portion des Signals vorgenommen.

Die Messung der Formanten diente der Untersuchung des Zusammenhangs zwischen Vokaldauer und Öffnungsgrad (vgl. 3.1, 5.3).¹³ Die Vokalformanten wurden in Spektren gemessen, die an einem bestimmten Zeitpunkt die Obertonfrequenzen (Abszisse) gegen deren Intensität bzw. deren Schalldruckpegel (Ordinate) auftragen und somit einen momentanen Querschnitt durch ein Spektrogramm eröffnen. Dabei wurde jeweils der durchschnittliche Energieverlauf des mittleren Drittels des Vokals berechnet, um so gut wie möglich allfällige Koartikulationserscheinungen auszumerzen. Abb. 2 zeigt das Spektrum des Vokals [i] im von

Frequenzband durchlässt, zeigt das Breitbandspektrogramm eine genauere Zeitauflösung und eignet sich somit besser für den Zweck der Segmentierung (vgl. Reetz 1999: 70-80).

Die hier abgebildeten Grafiken wurden mit dem Programm *Praat* erstellt (<http://www.fon.hum.uva.nl/praat/>).

¹³ Die Klangfarbe der Vokale kann am physikalischen Korrelat der 'Formanten' festgemacht werden (siehe Neppert ⁴1999: 129-135, Reetz 1999: 135-141). Darunter versteht man besonders energiereiche Obertonbereiche eines Klangs, die aufgrund der Beschaffenheit des Resonanzraums zustande kommen. Als Masswert für die fortlaufend nummerierten Formanten (F1, F2, usw.) wird in der Regel die Frequenz des Intensitätsmaximums in Hertz angegeben (bzw. die 'Formantmittenfrequenz'; vgl. Neppert 1999: 255).

Für die artikulatorische Interpretation dieser Werte gilt, dass der erste Formant umgekehrt proportional zum Öffnungsgrad eines Vokals ist (d.h. je tiefer der Vokal, desto höher F1), während der zweite Formant mit der Länge des Ansatzrohrs, d.h. mit den horizontalen Bewegungen der Zunge und der Lippenrundung korreliert (Vordervokale haben einen höheren F2-Wert als Hintervokale, gerundete Vorderzungenvokale haben einen niedrigeren F2-Wert als ungerundete).

der Sprecherin A geäußerten Wort [ˈziβə]; die Pfeile weisen auf den ersten Formanten bei 280 Hz bzw. auf den zweiten Formanten bei 2480 Hz.

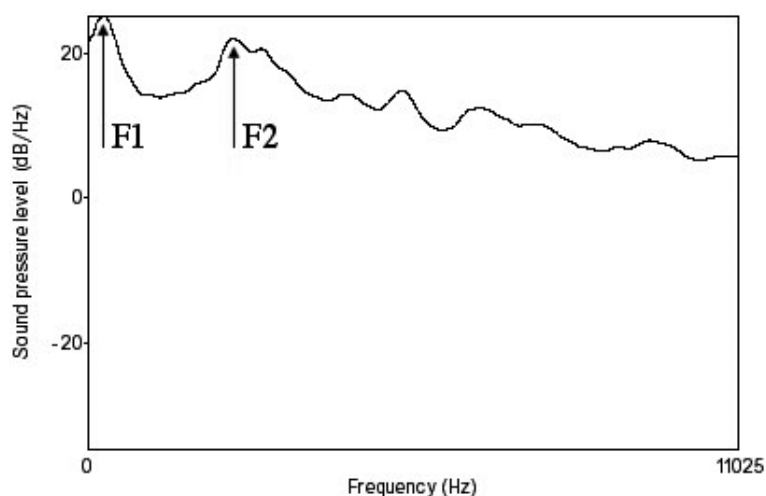


Abb. 2 Lage der ersten beiden Formanten im Spektrum des Vokals [i]

5 Ergebnisse

5.1 Die Dauer der Kurz- und Langvokale

Das durchschnittliche Verhältnis eines Kurzvokals zu einem Langvokal beträgt im Gesamtkorpus 0.56 und weist somit einen für den Quantitätskontrast plausiblen Wert auf, wenn man ihn mit den in 2.3 erwähnten Sprachen (Finnisch 0.44, Dänisch 0.50, Serbokroatisch 0.67) vergleicht. Stellt man nun einen Vergleich mit anderen deutschen Dialekten an (Spiekermann 2000: 221-222), so nimmt die Zürcher Mundart ebenfalls eine mittlere Stellung ein zwischen den Extremwerten des Rheinfränkischen (0.39) und des Mittelbairischen (0.82); vor allem aber liegt unser V/V:-Verhältnis relativ nahe bei den 0.62 des hochalemannischen Dialekts von Neuenburg (D).

Bezüglich der internen Differenzierung des Korpus ist zuerst einmal eine gewisse Streuung der Durchschnittswerte unter den drei Informanten zu vermerken: so zeichnet sich die Sprecherin A mit 0.47 durch ein niedrigeres V/V:-Verhältnis aus als die beiden männlichen Sprecher B und C, die beide einen Wert von 0.64 aufweisen. Der Sprechstil der Sprecherin A erscheint aufgrund des Höreindrucks als sorgfältiger, d.h. langsamer und eher zur Hyperartikulation neigend; man könnte demnach die - noch zu verifizierende - Hypothese aufstellen, dass beim *Lento*-Sprechstil das Dauerverhältnis zwischen Lang- und Kurzvokalen deutlicher ausfällt als beim *Allegro*-Stil.

Die Tab. 1 listet für die neun im Gesamtkorpus untersuchten Klangfarben die Mittelwerte der Dauer von Kurz- und Langvokalen (in Klammer die jeweilige Standardabweichung) sowie das daraus resultierende V/V:-Verhältnis auf:

| | i | y | u | e | ø | o | ɛ | æ | ɒ | Mittelwert |
|------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|------------|
| V | 54 (±12) | 95 (±19) | 96 (±23) | 110 (±26) | 87 (±15) | 81 (±15) | 102 (±22) | 131 (±29) | 109 (±7) | 96 |
| V: | 159 (±39) | 189 (±51) | 178 (±50) | 185 (±57) | 167 (±31) | 128 (±38) | 161 (±32) | 177 (±48) | 192 (±35) | 171 |
| V/V: | 0.34 | 0.50 | 0.50 | 0.60 | 0.52 | 0.63 | 0.63 | 0.74 | 0.57 | 0.56 |

Tab. 1 Mittelwert und Standardabweichung der Vokaldauer (in Millisekunden) nach Klangfarbe

Das Dauerverhältnis zwischen Kurz- und Langvokal liegt bei den meisten Klangfarben zwischen 0.50 und 0.63 und entfernt sich somit nicht sehr vom Mittelwert 0.56. Hingegen reissen die Werte von [æ] (0.74) und [i] (0.34) nach oben bzw. nach unten aus. Im ersten Fall wird das V/V:-Verhältnis vor allem durch die lange Dauer des Kurzvokals bestimmt - man beachte den Mittelwert von 131 ms, der länger ausfällt als derjenige des kürzesten Langvokals /o:/ (128 ms), und die ebenfalls relativ hohe Standardabweichung von ±29 ms: hier scheint tatsächlich die quasi-präpausale Stellung der Interjektion *gäll* im Satz (25) den Vokal extrinsisch gedehnt zu haben (vgl. 4.1). Beim zweiten Ausreisser [i] summiert sich vermutlich die intrinsisch kurze Dauer des Hochvokals mit der extrinsischen Kürzung aufgrund des Artikulationsorts des folgenden bilabialen Plosivs /b/ (vgl. 2.3). Ansonsten ist die Wirkung der intrinsischen Vokaldauer (vgl. 2.3) in der Tab. 1 zumindest bei den Tiefvokalen nicht klar ersichtlich: die Realisierungen von /ɒ:/ erreichen zwar die längste Dauer, aber umgekehrt sind die Realisierungen von /æ:/ kürzer als diejenigen der Hochvokale /y:/ und /u:/.

Im allgemeinen werden auch die Erwartungen bezüglich der extrinsischen Vokaldauer nicht erfüllt, zumindest was die Beeinflussung durch den nachfolgenden Konsonanten betrifft. Gemäss Voraussage sollten Vokale z.B. vor stimmhaften Konsonanten eine längere Dauer aufweisen als vor stimmlosen Konsonanten (vgl. 2.3) - eine Annahme, die durch die Daten der Tab. 2 nur zum Teil erhärtet wird.

| __ C [-stimmhaft] | | | | Mittelwert | __ C [+stimmhaft] | | | | | Mittelwert |
|-------------------|-----|-----|-----|---------------|-------------------|-----|-----|-----|-----|--------------|
| i | e | ø | u | | o | y | æ | ɛ | ɒ | |
| 54 | 110 | 87 | 96 | 86.75 | 81 | 95 | 131 | 102 | 109 | 103.6 |
| i: | e: | ø: | u: | | o: | y: | æ: | ɛ: | ɒ: | |
| 159 | 185 | 167 | 178 | 172.25 | 128 | 189 | 177 | 161 | 192 | 169.4 |

Tab. 2 Vokaldauer (in Millisekunden) vor stimmlosen und stimmhaften Konsonanten

Die Vokale in der linken Hälfte der Tabelle stehen vor den stimmlosen Obstruenten [b], [t], [g̊ k], [x], während die Vokale auf der rechten Seite von den Sonoranten von [l], [r] und [n] gefolgt werden. Die Annahme, dass Vokale vor stimmhaften Konsonanten länger seien, trifft in unserem Fall also nur für die Kurzvokale zu, nicht aber für die Langvokale.

Ähnlich verhält es sich mit der extrinsischen Vokaldauer in Bezug auf die phonotaktische und prosodische Struktur des Wortes. Die Tab. 3 stellt die Vokaldauer der Einsilbler derjenigen der Zweisilbler gegenüber:

| C(C)VC | | | | Mittelwert | 'CV.CV | | | | | Mittelwert |
|--------|-----|-----|-----|---------------|--------|-----|-----|-----|-----|-------------|
| e | ɛ | æ | u | | i | y | ø | o | ɒ | |
| 110 | 102 | 131 | 96 | 109.75 | 54 | 95 | 87 | 81 | 109 | 85.2 |
| e: | ɛ: | æ: | u: | | i: | y: | ø: | o: | ɒ: | |
| 185 | 161 | 177 | 178 | 175.25 | 159 | 189 | 167 | 128 | 192 | 167 |

Tab. 3 Vokaldauer (in Millisekunden) in Einsilblern und Zweisilblern

Die Vokale in der linken Hälfte der Tab. 3 kommen in einsilbigen Wörtern vom Typus *Bett*, *Herr*, *gäll* und *Bruch* vor, während die Vokale auf der rechten Seite in zweisilbigen Wörtern wie *sibe*, *Füli*, *Böge*, *hole* und *Mane* erscheinen. Auch hier kann von extrinsischer Vokaldauer (im Sinne einer kompensatorischen Dehnung bei den Einsilblern) nur bei den Kurzvokalen die Rede sein, während sich die entsprechenden Werte der Langvokale kaum unterscheiden. Immerhin fällt das V/V:-Dauerverhältnis bei den Einsilblern mit 0.62 weniger ausgeprägt aus als bei den Zweisilblern (0.51); aufgrund von an anderen Sprachen gemachten Beobachtungen (vgl. 2.3) kann man dieses Ergebnis dahingehend interpretieren, dass die Vokale in CVC-Wörtern weniger komprimierbar sind.¹⁴

Die Gesamtwertung der Dauerverhältnisse von Kurz- und Langvokalen in unserem Korpus ergibt, dass die bekannten Faktoren der intrinsischen und extrinsischen Vokaldauer sich nur zum Teil als relevant erweisen, indem sie gewisse Abweichungen vom Mittelwert des V/V:-Verhältnisses zu erklären vermögen. Hingegen haben bis auf eine Ausnahme alle Kurzvokale eine deutlich geringere Dauer als die entsprechenden Langvokale und die Abweichungen vom Mittelwert 0.56 fallen relativ gering aus: diese Tatsache belegt eindeutig, dass der wesentliche Grund für die Dauerunterschiede in der distinktiven Funktion der Quantitätsopposition liegt und somit letztendlich auf der Lautabsicht der Sprecher beruht.

5.2 Die Dauer der Konsonanten nach Lang- und Kurzvokalen

Im Zusammenhang mit der Vokalquantität stellt sich die Frage, inwieweit die unterschiedliche Vokaldauer durch Unterschiede in der Dauer der Konsonanten kompensiert wird. In ihren Erhebungen zu einem norditalienischen Dialekt der Emilia-Romagna stellen z.B. Uguzzoni/Busà (1995: 391-392) fest, dass bei zweisilbigen Wörtern keine Unterschiede vorliegen, während die Werte bei Einsilblern je nach Sprecher z.T. beträchtlich variieren.

Die Tab. 4 zeigt die jeweiligen Dauerverhältnisse zwischen den Konsonanten, die auf die Kurz- und Langvokale folgen.

¹⁴ Eine entgegengesetzte Tendenz stellt Gredig (2000: 61) im Bündnerromanischen von Scuol fest, wo die prozentuale Dauer der Kurzvokale im Vergleich zu den Langvokalen bei den Einsilblern 0.45, bei den Zweisilblern 0.52 beträgt.

| | | | | | |
|-------------------|-------------------------|------|-------------|------------------------|------|
| CVC | <i>gäll vs. gääl</i> | 0.97 | CVCV | <i>Böge vs. Böögge</i> | 0.66 |
| | <i>Herr vs. hèèr</i> | 0.97 | | <i>sibe vs. siibe</i> | 0.87 |
| | <i>Bett vs. Beet</i> | 1.06 | | <i>Füli vs. Fööli</i> | 0.89 |
| | <i>Bruch vs. Bruuch</i> | 1.23 | | <i>hole vs. hoole</i> | 1.01 |
| | | | | <i>Mane vs. maane</i> | 1.04 |
| Mittelwert | | 1.06 | | | 0.89 |

Tab. 4 Verhältnis der Konsonantendauer (Mittelwerte) nach Kurz- und Langvokalen in Ein- und Zweisilblern

Geht man von der Nullhypothese und somit von der Annahme aus, dass die Konsonantendauer nicht durch die Vokalquantität beeinflusst wird, so sollten die entsprechenden Werte in jedem Fall möglichst 1.00 erreichen. Nimmt man hingegen an, dass die Konsonantendauer ein mögliches Korrelat der Quantitätsopposition darstellt, so sollte dieses Verhältnis vor allem bei Einsilblern den Wert 1.00 deutlich überschreiten. Wie man der Tabelle entnehmen kann, trifft diese Annahme nur bis zu einem gewissen Grad für den Frikativ /x/ zu, während bei den anderen Konsonanten in CVC-Wörtern die Dauerverhältnisse sehr nahe bei 1.00 liegen.

Auch bei den CVCV-Wörtern besteht kein wesentlicher Unterschied in der Konsonantendauer nach Kurz bzw. Langvokal. Absolut einleuchtend ist der Sonderfall von *Böge vs. Böögge* (0.66), da es sich hier um ein Pseudo-Minimalpaar handelt, bei welchem neben dem Kontrast in der Vokalquantität auch eine phonemische Opposition zwischen /g̃/ ~ /k/ besteht; in der Tat liegt dieses Verhältnis im Bereich der Werte, die Willi (1996: 140-141) für intersonore *lenis*- und *fortis*-Plosive erhoben hat.¹⁵ Wenn man sinnvollerweise dieses Pseudo-Minimalpaar bei der Berechnung des Mittelwertes nicht berücksichtigt, so ergibt sich für die zweisilbigen Wörter ein Verhältnis der Konsonantendauer nach Kurz- und Langvokal von 0.95 (anstatt von 0.89 wie im Feld rechts unten der Tab. 4).

Demzufolge treten im Zürichdeutschen weder bei Einsilblern noch bei Zweisilblern bemerkenswerte Kompensationserscheinungen in der Konsonantendauer auf, weshalb letztere als mögliches akustisches Korrelat der vokalischen Quantitätsopposition ausser Betracht fallen muss.

5.3 Die Klangfarbe der kurzen und langen Hochvokale

Zu klären bleibt nun die Frage, inwieweit die Längenunterschiede der Vokale auch von klar wahrnehmbaren Veränderungen der Klangfarbe begleitet werden. Ein solcher Zusammenhang besteht unbestrittenermassen im Standarddeutschen (vgl. 2.1) und wird von Keller (1961) auch für die Mundart der Stadt Zürich postuliert (vgl. 3.1). Die Frage ist nun, ob die

¹⁵ Die Konsonantendauer wird hier als akustisches Korrelat einer phonemischen Opposition zwischen *fortes* und *lenes* verstanden, im Gegensatz etwa zur Analyse der Thurgauer Mundart durch Kraehenmann (2001), die von einem zugrundeliegenden Kontrast zwischen Geminaten und einfachen Konsonanten ausgeht. Eine dritte Lösung schlägt Ham (2001: 52) für das Berndeutsche vor in Form einer Dreiteilung der Obstruenten in *lenes*, einfache und geminierte *fortes*.

natürliche Korrelation zwischen Länge und Gespanntheit in unserem Dialekt ein Ausmass erreicht, das wie im Deutschen die Annahme von zwei verschiedenen Öffnungsgraden - halbhoch und hoch - für Kurz- und Langvokale rechtfertigen würde, oder ob allfällige Unterschiede in der Klangfarbe eher im subphonematischen Bereich anzusiedeln sind, wie dies etwa im von Uguzzoni/Busà (1995: 392-393) untersuchten norditalienischen Dialekt der Fall zu sein scheint.

Zur akustischen Überprüfung des Zusammenhangs zwischen Vokaldauer und Öffnungsgrad wurden je neun Realisierungen (drei pro Sprecher) der sechs Vokalphoneme /i i: y y: u u:/ untersucht, wobei jeweils der erste und der zweite Formant gemessen wurde (vgl. 4.3). Die tendenzielle Verteilung dieser Vokalrealisierungen wird in Abb. 3 dargestellt, in welcher die F1-Werte auf der Ordinate und die F2-Werte auf der Abszisse aufgetragen werden; um eine Korrespondenz zur üblicherweise verwendeten und artikulatorisch motivierten Darstellung der Vokale zu erreichen, wurde die Skalierung umgekehrt, so dass sich der Ursprung des kartesischen Diagramms rechts oben befindet. Die Lage und Form der Ellipsen für die verschiedenen Vokale wird durch die Mittelwerte sowie die Standardabweichung der jeweiligen F1 und F2-Werte bestimmt.

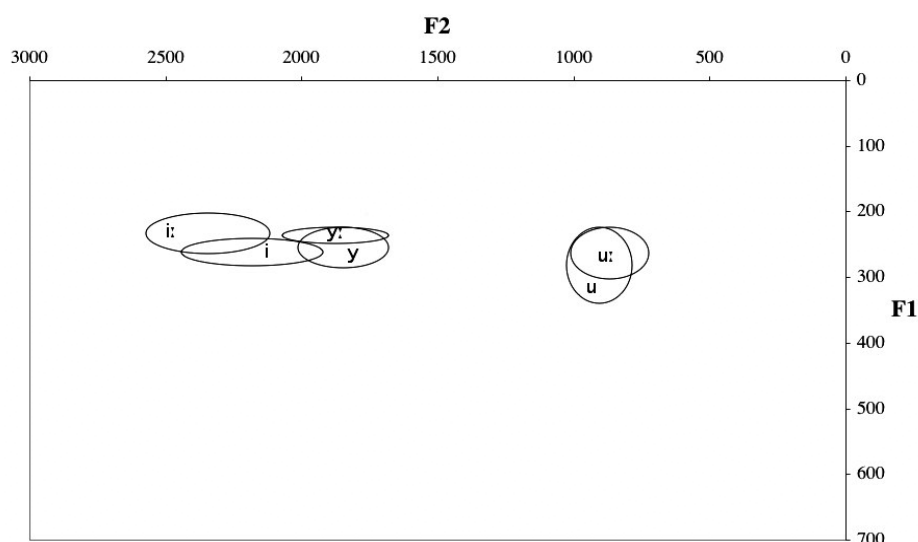


Abb. 3 Tendenzielle Streuung der kurzen und langen Hochvokale im Zürichdeutschen

Vergleicht man nun die drei durch die Quantitätsopposition unterschiedenen Vokale, so zeigt sich, dass die Kurzvokale tatsächlich eher tiefer liegen oder zumindest - wie im Fall von [y] und [u] - eine etwas grössere Streuung nach unten bzw. gegen das Zentrum der Vokalfläche aufweisen, während auf der anderen Seite die Ellipsen der Langvokale minim höher gelegen sind und sich eher peripher ausdehnen. Ein gewisser Zusammenhang zwischen Länge und Gespanntheit ist also nicht von der Hand zu weisen. Umgekehrt muss betont werden, dass die Ellipsen von Kurz- und Langvokalen sich in allen drei Fällen überschneiden und dass die Überschneidung bei [y] und [u] sogar ziemlich deutlich ausfällt. Allein schon dieser Umstand legt den Schluss nahe, dass die entsprechenden Kurz- und Langvokale im wesentlichen den gleichen Öffnungsgrad haben und sich somit zwar sehr wohl in der Dauer, nicht aber in der Klangfarbe unterscheiden.

Dieser Sachverhalt wird noch klarer, wenn man die F1- und F2-Mittelwerte der untersuchten Vokale den entsprechenden Kennwerten des Standarddeutschen gegenüberstellt. In der Tab. 4 werden drei Vokalkategorien unterschieden, nämlich lange Hochvokale, kurze Hochvokale und kurze Halbhochvokale; da sowohl im Zürichdeutschen als auch im Standarddeutschen drei Vokalreihen vorkommen (ungerundete Vordervokale, gerundete Vordervokale und Hintervokale), ergeben sich daraus neun Spalten. Die drei Zeilen der Tabelle beziehen sich auf zwei Untersuchungen zum Standarddeutschen, abgekürzt mit DE2 (Delattre 1981: 73) und DE1 (Rausch 1972; gerundete statistische Mittelwerte, zitiert nach Pétursson/Neppert ²1996: 139), sowie auf die eigenen Messungen zum Zürichdeutschen.

| | | i: | i | ɪ | y: | y | ʏ | u: | u | ʊ |
|-----|----|------|------|------|------|------|------|-----|-----|-----|
| DE2 | F1 | 300 | | 350 | 300 | | 350 | 300 | | 375 |
| | F2 | 2300 | | 2100 | 1750 | | 1600 | 825 | | 875 |
| DE1 | F1 | 250 | | 363 | 250 | | 400 | 250 | | 369 |
| | F2 | 2400 | | 2088 | 1563 | | 1440 | 688 | | 750 |
| ZH | F1 | 236 | 262 | | 240 | 262 | | 267 | 298 | |
| | F2 | 2342 | 2173 | | 1880 | 1849 | | 871 | 911 | |

Tab. 5 Mittelwerte für F1 und F2 (in Hertz) von Hochvokalen im Zürichdeutschen und im Standarddeutschen

Die Unterteilung in drei verschiedene Vokalkategorien wird durch die Mittelwerte der Formanten eindeutig gerechtfertigt: vor allem bezüglich des mit dem Öffnungsgrad korrelierenden ersten Formanten liegt die mittlere Kategorie der zürichdeutschen Kurzvokale /i y u/ (Mittelwerte: 262 Hz, 262 Hz, 198 Hz) viel näher bei den zürichdeutschen Langvokalen /i: y: u:/ (Mittelwerte: 236 Hz, 240 Hz, 267 Hz) als bei den ungespannten Kurzvokalen von DE 1 (Mittelwerte: 363 Hz, 400 Hz, 369 Hz). Bezüglich des mit der horizontalen Zungenlage korrelierenden zweiten Formanten liegt das zürichdeutsche /i/ mit 2173 Hz in der Mitte zwischen den beiden anderen Kategorien, wenn man es mit den Mittelwerten für zürichdeutsches /i:/ (2342 Hz) und /ɪ/ (2088 Hz) im Korpus DE1 vergleicht. Das zürichdeutsche /y/ liegt mit einem mittleren F2-Wert von 1849 Hz sogar peripherer als der entsprechende Langvokal des Standarddeutschen, für welchen das DE1-Korpus einen F2-Mittelwert von 1563 Hz aufweist. Einzig das zürichdeutsche /u/ zeigt mit einem mittleren F2-Wert von 911 Hz eine etwas stärkere Zentralisierung als die entsprechenden deutschen Kurzvokale in DE1 (Mittelwert des zweiten Formanten: 750 Hz).

Angesichts dieser Datenlage kann man schliessen, dass der phonologische Längekontrast bei den Hochvokalen des Zürichdeutschen nicht durch nennenswerte Qualitätsunterschiede begleitet wird.

5.4 Silbenschnitt?

Bis hierhin hat sich die Vokaldauer als das einzige phonetische Korrelat der Quantitätsopposition herauskristallisiert, während die Dauerunterschiede der nachfolgenden Konsonanten ebenso wie die vokalische Klangfarbe vernachlässigt werden können. Ein weiteres oft mit der Quantität in Verbindung gebrachtes Korrelat muss an dieser Stelle

zumindest kurz angesprochen werden, nämlich der Intensitätsverlauf des Vokals im Kontakt zum nachfolgenden Konsonanten. Dieses Phänomen sollte man vor allem auch deshalb erwähnen, weil es in jüngster Zeit Gegenstand intensiver Forschung zum Deutschen und seinen Dialekten gewesen ist (Spiekermann 2000, Restle 2001, Auer/Gilles/Spiekermann 2002).

Es würde allerdings zu weit führen, die damit verbundenen Begriffe - 'scharfer' vs. 'sanfter' Silbenschnitt bzw. 'fester' vs. 'loser' Anschluss - ausführlich zu diskutieren, weil damit je nach Autor sehr verschiedene Aspekte einer breiten Thematik gemeint sind: einerseits kann man den 'Anschluss' als phonetisches Korrelat einer phonologischen Quantitätsopposition betrachten (etwa im Sinne des oben erwähnten vokalischen Intensitätsverlaufs), während andererseits der 'Silbenschnitt' selbst als phonologische Korrelation verstanden wird (im Sinne einer übergeordneten prosodischen Kategorie), von der eine Reihe phonologischer und phonetischer Phänomene wie die Vokalquantität und -qualität, der Intensitätsverlauf oder sogar die Ambisilbizität des folgenden Konsonanten abgeleitet werden kann.¹⁶

Von direktem Interesse im Zusammenhang mit der hier vorgestellten Untersuchung ist die umfangreiche Arbeit von Spiekermann (2000), in welcher ein operationalisierbares akustisches Korrelat des Silbenschnitts formuliert und an zahlreichen deutschen Dialekten getestet wird. Demnach stellen die Dauer und die damit verbundene Anzahl der Energie- bzw. Intensitätsmaxima in einem Vokal phonetische Korrelate sowohl der Quantität als auch des Silbenschnitts dar, während für den (scharfen) Silbenschnitt allein eigentlich nur zwei Korrelate in Frage kommen, nämlich ein relativ frühes Erreichen des Energiemaximums und ein tendenziell schnell steigend und abfallender (d.h. nicht anhaltender) Intensitätsverlauf.

Im Bezug auf die Mundart der Stadt Zürich ist vor allem die arealtypologische Schlussfolgerung interessant, wonach eigentliche Silbenschnittdialekte in Norddeutschland sowie z.T. im nieder- und mitteldeutschen Raum vorkommen; in Süddeutschland hingegen neigen vor allem die oberdeutschen Dialekte dazu, keinen Kontrast zwischen sanftem und scharfem Schnitt auszubilden (Spiekermann 2002: 197). Eine solche Arealtypologie legt die Vermutung nahe, dass im Zürichdeutschen die Quantität und nicht der Silbenschnitt die ausschlaggebende phonologische Eigenschaft bildet - eine Hypothese, die ja auch durch die relativ konstante Klangfarbe von Kurz- und Langvokalen (vgl. 5.3) gestützt wird.

Aufgrund dieser schwachen Plausibilität der Silbenschnittkorrelation wurde in der vorliegenden Untersuchung von einer quantitativen Analyse der Intensitätsverläufe abgesehen; einige wenige Stichproben erhärten aber die oben gemachte Annahme. Die Abbildungen 4 und 5 stellen die Spektrogramme sowie die Intensitätsverläufe der beiden von der Sprecherin A geäußerten Wörter *sibe* und *siibe* dar.

¹⁶ Für eine ausführliche Behandlung - aus sowohl theoretisch-systematischer als auch wissenschaftsgeschichtlicher Perspektive - der hier sehr verkürzt dargestellten Problematik sei auf die Darstellungen in Spiekermann (2000: 1-35) und Restle (2001: 1-89) hingewiesen.

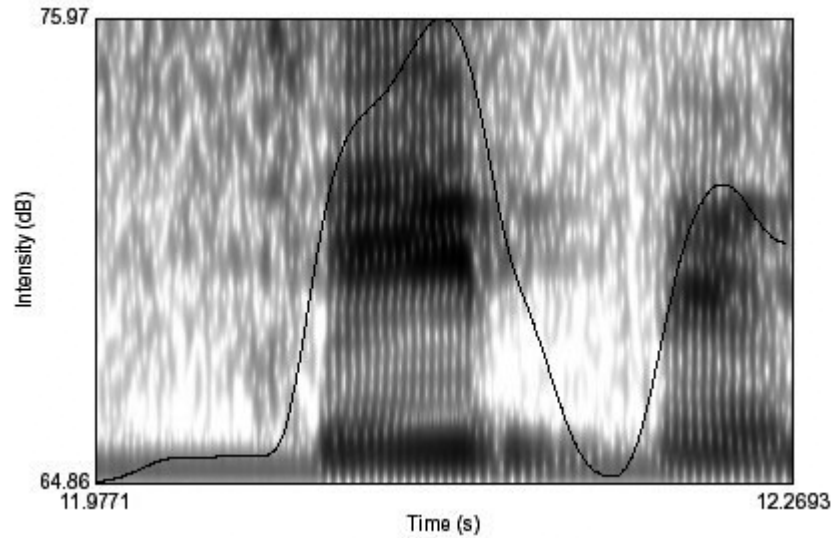


Abb. 4 Spektrogramm und Intensitätsverlauf des Wortes [ˈzi:bə]

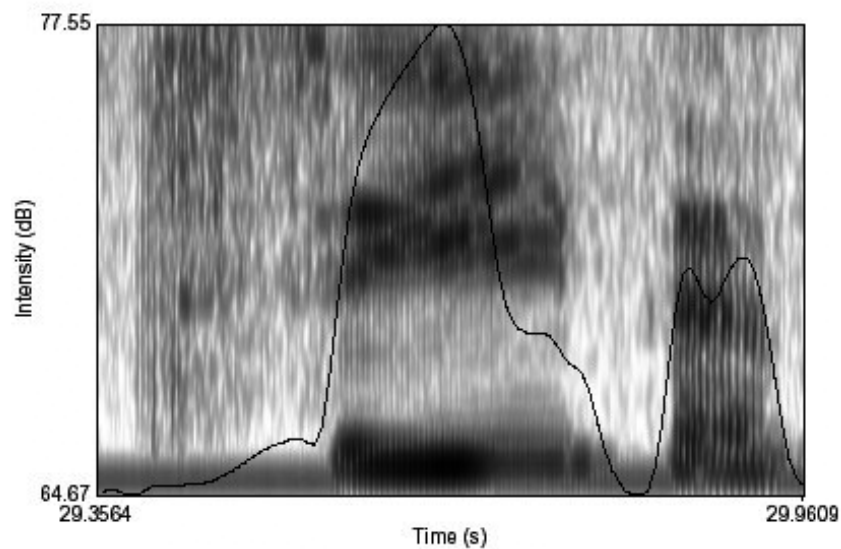


Abb. 5 Spektrogramm und Intensitätsverlauf des Wortes [ˈzi:bə]

Der Vergleich zwischen beiden Wörtern zeigt, dass der Kurzvokal seinen Energiehöhepunkt sehr spät (nach 51 Millisekunden bei einer Gesamtvokaldauer von 65 Millisekunden) erreicht, was dem typischen Intensitätsverlauf bei scharfem Silbenschnitt klar zuwiderläuft. Das Energiemaximum tritt hingegen beim Langvokal relativ früh ein, nämlich nach 91 ms bei einer Gesamtvokaldauer von 197 ms; von einem eigentlichen Halten der Intensität kann hier aber nicht die Rede sein, selbst wenn auf einem eher tiefen Niveau das Abfallen der Energie kurz verlangsamt wird. Um fundierte Aussagen zu diesem Aspekt der Kurz- und Langvokale im Zürichdeutschen machen zu können, müssten weitere z.T. sehr aufwändige Analysen gemacht werden; aufgrund der arealtypologischen Lokalisierung und der Sichtung einzelner Stichproben ist allerdings kaum damit zu rechnen, die typischen akustischen Korrelate einer Silbenschnittsprache vorzufinden.

6 Schlussbemerkung

An dieser Stelle muss betont werden, dass die hier gemachten Aussagen vorläufig nur für das Zürichdeutsche Geltung haben. Es wäre interessant zu sehen, wie sich die Quantität in den Vokalsystemen anderer schweizerdeutscher Dialekte phonologisch manifestiert und welche phonetischen Korrelate dabei zum Vorschein kommen.

Die vorliegende Untersuchung zeigt hingegen klar, dass die Mundart der Stadt Zürich die phonologische Eigenschaft der Vokalquantität aufweist. Dieses prosodische Merkmal prägt das Vokalsystem in grundlegender Art und Weise, findet man doch für neun von zehn Klangfarben Minimalpaare.

Was die phonetische Realisierung der Quantitätsopposition betrifft, so erweist sich die Vokaldauer als solides akustisches Korrelat. Das durchschnittliche Dauerverhältnis zwischen Kurz- und Langvokalen von 0.56 darf als typisch für eine Quantitätssprache bezeichnet werden; die relativ gering ausfallenden Abweichungen von diesem Mittelwert können im wesentlichen mit den bekannten Faktoren der intrinsischen und extrinsischen Vokaldauer erklärt werden. Die Vokalquantität bleibt hingegen ohne Folge auf die Dauer der folgenden Konsonanten. Angesichts widersprüchlicher Aussagen in der dialektologischen Literatur ist zudem als wichtiges Ergebnis festzuhalten, dass die Vokaldauer auch nicht mit nennenswerten Veränderungen der Klangfarbe einhergeht; das Zürichdeutsche hat keine phonologischen Halbhochvokale und zeichnet sich somit durch ein vierstufiges Vokalsystem aus. Was schliesslich den Intensitätsverlauf der Kurz- und Langvokale betrifft, so darf man aufgrund der vorliegenden Indizen davon ausgehen, dass die typischen Korrelate des Silbenschnitts fehlen.

Abschliessend kann somit festgestellt werden, dass die Mundart der Stadt Zürich eine typische Quantitätssprache darstellt, welche dieses prosodische Merkmal ausschliesslich durch die unterschiedliche Dauer von Kurz- und Langvokalen zum Ausdruck bringt.

Literaturangaben

- Auer, Peter/Couper-Kuhlen, Elizabeth/Müller, Frank (1999): *Language in time. The Rhythm and Tempo of Spoken Interaction*. New York etc.
- Auer, Peter/Gilles, Peter/Spiekermann, Helmut (eds.) (2002): *Silbenschnitt und Tonakzente*. Tübingen.
- Becker, Thomas (1998): *Das Vokalsystem der deutschen Standardsprache*. Bern.
- Bosoni, Giorgio (1995): "Dialettologia lombarda: un esempio di approccio strumentale allo studio delle opposizioni di quantità vocalica in sillaba tonica". *Studi Italiani di Linguistica Teorica e Applicata* 24: 345-364.
- Christen, Helen (2002): "'Swissdate' - Aspekte dialektalen Sprechens im Schweizer Privatfernsehen". In: Wiesinger, Peter (ed.): *Akten des X. Internationalen Germanistenkongresses Wien 2000 »Zeitenwende - Die Germanistik auf dem Weg vom 20. ins 21. Jahrhundert«*. Bern: 217-225.
- Dieth, Eugen (1938, ²1986): *Schwyzertütschi Dialäktschrift*. 2. Auflage herausgegeben von Christian Schmid-Cadalbert. Aarau.

- Dieth, Eugen (1950): *Vademekum der Phonetik*. Bern.
- Fox, Anthony (2000): *Prosodic Features and Prosodic Structure*. Oxford.
- Gredig, Stephanie (2000): *Fonetica e fonologia del vocalismo tonico nel romancio di Scuol (Bassa Engadina) con uno studio sperimentale delle durate vocaliche*. Lizentiatsarbeit, Universität Zürich.
- Haas, Walter (1978): *Sprachwandel und Sprachgeographie. Untersuchungen zur Struktur der Dialektverschiedenheit am Beispiele der Schweizerdeutschen Vokalsysteme*. Wiesbaden.
- Ham, William H. (2001): *Phonetic and phonological aspects of geminate timing*. New York/London.
- Hofer, Lorenz (1997): *Sprachwandel im städtischen Dialektrepertoire. Eine variations-linguistische Untersuchung am Beispiel des Baseldeutschen*. Basel/Tübingen.
- Hotzenköcherle, Rudolf (1986): "Aspekte und Probleme der Vokalquantität im Schweizerdeutschen". In: Schläpfer, Robert/Trüb, Rudolf (eds.): *Dialektstrukturen im Wandel: Gesammelte Aufsätze zur Dialektologie der deutschen Schweiz und der Walsergebiete Oberitaliens*. Aarau: 319-333.
- Keller, Rudolf Ernst (1961): *German dialects*. Manchester.
- Kohler, Klaus (²1995): *Einführung in die Phonetik des Deutschen*. Berlin.
- Kraehenmann, Astrid (2001): "Swiss German stops: geminates all over the word". *Phonology* 18: 109-145.
- Kraehenmann, Astrid (2003): "Swiss German vowel length through time". In: Fikkert, Paula/Jacobs, Haike (eds.): *Development in Prosodic Systems*. Berlin: 283-314.
- Ladefoged, Peter/Maddieson, Ian (1996): *The Sounds of the World's Languages*. Oxford.
- Laver, John (1994): *Principles of Phonetics*. Cambridge.
- Lehiste, Ilse (1970): *Suprasegmentals*. Cambridge.
- Maddieson, Ian (1984): *Patterns of Sounds*. Cambridge.
- Martinet, André (1962): *A functional view of language*. Oxford. (französische Übersetzung 1969. *Langue et fonction*. Paris).
- Moulton, William G. (1973): "Vowel systems with five heights". In: Scholler, Harald/Reidy, John (eds.): *Lexicography and dialect geography*. Wiesbaden: 187-194.
- Neppert, Joachim M. H. (⁴1999): *Elemente einer akustischen Phonetik*. Hamburg.
- Pétursson, Magnús/Neppert, Joachim M. H. (²1996): *Elementarbuch der Phonetik*. Hamburg.
- Ramers, Karl Heinz (1988): *Vokalquantität und -qualität im Deutschen*. Tübingen.
- Rausch, Arsen (1972): "Untersuchungen zur Vokalartikulation im Deutschen". In: Kelz, Heinrich/Rausch, Arsen: *Beiträge zur Phonetik*. Hamburg: 35-82. (= *IPK-Forschungsberichte* 30).
- Reetz, Henning (1999): *Artikulatorische und akustische Phonetik*. Trier.
- Restle, David (2001): *Silbenschnitt - Quantität - Kopplung. Zur Geschichte, Charakterisierung und Repräsentation der Anschlussprosodie unter dem Blickwinkel einer Oszillationssilbentheorie*. Diss. Universität München.
- Schmid, Stephan (1999a): *Fonetica e fonologia dell'italiano*. Torino.
- Schmid, Stephan (1999b): "Per un'analisi del vocalismo italo-romanzo in chiave tipologica". In: Benincà, Paola/Mioni, Alberto/Vanelli, Laura (eds.): *Fonologia e morfologia dell'italiano e dei dialetti d'Italia*. Roma: 249-267.

- Schobinger, Viktor (2001): *Säit me soo oder andersch?* Zürich.
- Schwab, Sandra/Keller, Eric/Zellner, Brigitte/Connan, Pierre-Yves/Siebenhaar, Beat (1998): *Conventions de segmentation pour la construction de diphones*. LAIP, Université de Lausanne.
- SDS = *Sprachatlas der deutschen Schweiz* (hrsg. von Rudolf Hotzenköcherle). Vol. I *Lautgeographie: Vokalqualität* (1962). Vol. II *Lautgeographie: Vokalquantität, Konsonantismus* (1965). Bern.
- Sorianello, Patrizia (1997): "Parlato letto e parlato spontaneo. Indici prosodici a confronto". In: Cutugno, Francesco (ed.): *Fonetica e fonologia degli stili dell'italiano parlato*. Roma: 89-110.
- Spiekermann, Helmut (2000): *Silbenschnitt in deutschen Dialekten*. Tübingen.
- Spiekermann, Helmut (2002): "Ein akustisches Korrelat des Silbenschnitts: Formen des Intensitätsverlaufs in Silbenschnitt- und Tonakzentsprachen". In: Auer, P./Gilles, P./Spiekermann, H. (eds.): 181-199.
- Trautmüller, Hartmut (1982): "Vokalismus in der niederösterreichischen Mundart". *Zeitschrift für Dialektologie und Linguistik* 2: 289-333.
- Willi, Urs (1996): *Die segmentale Dauer als phonetischer Parameter von 'fortis' und 'lenis' bei Plosiven im Zürichdeutschen*. Stuttgart.
- Uguzzoni, Arianna/Busà, Maria Grazia (1996): "Acoustic correlates of vowel quantity contrasts in an Italian dialect". *Proceedings of the International Congress of the Phonetic Sciences*. Vol. III. Stockholm: 390-393.