Erschienen in: Belz, Malte/ Fuchs, Susanne/ Jannedy, Stefanie/ Mooshammer, Christine/ Rasskazova, Oksana/ Żygis, Marzena (Hrsg.) : Proceedings of the conference on phonetics & phonology in German-speaking countries (P&P 13). - Berlin: Humboldt-Universität zu Berlin, 2018. S. 77-80

Vergleich der Quantität, Qualität und Dynamik in den deutschen <a>-Lauten

Sandra Hansen-Morath, Anja Geumann, Renate Raffelsiefen

Institut für Deutsche Sprache, Mannheim

hansen@ids-mannheim.de, geumann@ids-mannheim.de, raffelsiefen@ids-mannheim.de

Abstract

In diesem Beitrag werden drei quantitative Studien vorgestellt, mit deren Hilfe untersucht wird, ob neben dem robusten Längenunterschied auch Qualitätsunterschiede für die deutschen <a>-Laute vorhanden sind (z.B. <Saat> versus <satt>). Auf Basis von ausgewählten Korpora und instrumental-phonetischen Messungen kann dieser Zusammenhang bestätigt werden. Zudem zeigen sich signifikante Unterschiede in den dynamischen Verläufen der beiden Vokale.

Einleitung

Minimalpaare des Deutschen wie *Schrot – Schrott*, *Ruhm – Rum*, *Saat – satt* sind durch eine Reihe korrelierender phonetischer Unterschiede hinsichtlich der Vokalqualität, Vokaldauer und Silbenstruktur gekennzeichnet. Phonologische Studien stützen die Annahme einer einheitlichen binären Opposition: sowohl das Neutralisationsverhalten wie auch die Evidenz bezüglich Paradigmenuniformität indizieren eine durchgehende Qualitätsopposition (Raffelsiefen 2016). Dieses Ergebnis steht im Einklang mit artikulatorischen Studien, die Qualitätsunterschiede für alle Paare zeigen (Bohn et al. 1992, Hoole/Mooshammer 2002), nicht aber mit Perzeptionsstudien, in denen für tiefe Vokale keine Qualitätsunterschiede festgestellt wurden (Weiss 1976).

In dem vorliegenden Beitrag werden quantitative Analysen vorgestellt, mit deren Hilfe Qualitätsunterschiede für die deutschen <a>-Laute nachgewiesen werden können. Hierzu werden <a>-Laute in hauptbetonten Silben auf Basis des robusten Längenunterschieds in zwei Klassen unterteilt (im Folgenden: <Saat>-Typ-Vokal vs. <satt>-Typ-Vokal). Es soll der Frage nachgegangen werden, ob die auf diese Weise unterschiedenen Vokalklassen korrelierende Qualitätsunterschiede, bzw. Unterschiede in den dynamischen Verläufen zeigen.

Datengrundlage

Als Datengrundlage für die Studien dienen verschiedene Korpora: das Kiel Korpus (Kohler 1992), das „Deutsch heute“ Korpus (Brinckmann et al. 2008), das Oldenburger Logatome Speech Corpus (Wesker et al. 2005) und ein projektinternes Korpus, das Nonsens-Minimalpaare einschließt, die von 29 (14m/15f) standardnahen Sprechern realisiert wurden. Die Untersuchung ist auf Wörter beschränkt, in denen <a>-Laute in hauptbetonten Silben erscheinen. Es wurde überprüft, dass in allen Fällen <Saat>-Typ-Vokale und <satt>-Typ-Vokale aufgrund von Längenunterschieden in den Annotationen angemessen unterschieden waren (vgl. hierzu auch die obere Grafik in Abb. 1).

Akustische Analyse

Die Formantwerte (F1, F2, F3) wurden mit PRAAT (Boersma et al. 2016) automatisch extrahiert. Dazu wurde der Burg-Algorithmus mit den Standardeinstellungen verwendet (5 Formanten im Suchraum bis 5500 Hz für Frauenstimmen und bis 5000 Hz für Männerstimmen). Die Messzeitpunkte liegen für unsere Auswertung bei 25%, 50% und 75% der akustischen Vokaldauer.

Neben den Formantwerten (von Hertz in Bark transformiert) werden außerdem Maße für den dynamischen Verlauf der jeweiligen Vokale errechnet. Es werden die Euklidischen Distanzen (in Bark) der F1xF2 Trajektorien von 25% bis 75% der Vokaldauer berechnet (Erläuterung (1)) sowie die Steigung (= slope) innerhalb der Formantverläufe F1 und F2 (vgl. Erläuterungen (2) und (3)).

(1) Euklidische Distanz zwischen 25%-75%: $\sqrt{((F1_{25\%} - F1_{75\%})^2) + ((F2_{25\%} - F2_{75\%})^2)}$.

(2) F2 slope: $F2_{75\%} - F2_{25\%} / 0.5 *$ Vokaldauer.

(3) F1 slope: $F1_{75\%} - F1_{25\%} / 0.5 *$ Vokaldauer.

Außerdem wird ein weiteres (nicht dynamisches) Maß erhoben, das die Peripheralität operationalisiert: ein zweites Euklidisches Distanzmaß (im Folgenden: Euklidische DistanzPER genannt), berechnet aus der Distanz der F1 Werte und F2 Werte in der Vokalmitte zu den F1 Mittelwerten und F2 Mittelwerten (in Bark) aller Männer bzw. Frauen für das jeweilige Korpus. Somit ist die Berechnung der Mitte jeweils spezifisch für ein bestimmtes (Sub)korpus (z.B. die 26 SprecherInnen des Kielkorpus).

Es geht uns insbesondere darum zu untersuchen, ob für die Unterscheidung der deutschen <a>-Laute neben als robust vorausgesetzten Daueruntersschieden weitere qualitative Unterschiede in F1, F2, F3 (in der Vokalmitte) und bezüglich der Peripheralität des Vokals sowie darüber hinaus dynamische Unterschiede in der Euklidischen Distanz und in F1 slope und F2 slope eine Rolle spielen.

Statistische Analyse

Um zu überprüfen, ob Qualitätsunterschiede und Unterschiede in den dynamischen Verläufen zwischen den beiden Vokalklassen (<Saat>-Typ-Vokal vs. <satt>-Typ-Vokal) vorliegen, werden mehrere gemischte Modelle berechnet, die es erlauben, die interindividuelle SprecherInnenvariation als Zufallsfaktor zu kontrollieren.

Im Rahmen der ersten Studie wird in mehreren gemischten linearen Modellen (Bates et al. 2017) der Einfluss des Vokaltyps (als Prädiktorvariable) auf die folgenden Kriteriumsvariablen getestet: ‚F1', ‚F2', ‚F3', ‚Euklidische Distanz', ‚Euklidische DistanzPER', ‚F1 slope' und ‚F2 slope' sowie zu Kontrollzwecken auch ‚Vokaldauer'. Um weitere Einflüsse kontrollieren zu können, werden außerdem das Geschlecht der SprecherInnen, der nachfolgende Laut im betreffenden Wort und die Information, ob es sich bei dem betreffenden Wort um ein sinnhaftes oder nicht sinnhaftes Wort handelt, als Prädiktorvariablen in das statistische Modell aufgenommen.

In der zweiten Studie werden mit Hilfe von logistischen Regressionsmodellen (Dobson 1990, Hastie/Pregibon 1992, McCullagh/Nelder 1989) die Einflüsse der Faktoren ‚F1', ‚F2', ‚F3', ‚Euklidische Distanz', ‚Euklidische DistanzPER' sowie ‚F1 slope', ‚F2 slope' und - zu Kontrollzwecken - ‚Vokaldauer' (als Prädiktorvariablen) auf den Vokaltyp (als Kriteriumsvariable) überprüft und hierarchisiert.

Die dritte Studie befasst sich zusätzlich mit regionalen Einflüssen auf die Vokalqualität und den dynamischen Verläufen.

Ergebnisse

Um zunächst regional bedingte Variation auszuschließen, werden in der ersten Studie nur folgende SprecherInnen berücksichtigt: SprecherInnen aus dem Deutsch heute Korpus aus Nord-West Deutschland und dem Standard zugeordnete SprecherInnen aus dem Oldenburger Logatome Speech Corpus. In beiden Korpora können signifikante Einflüsse des Vokaltyps auf F1 ($t=6.65$; $p<.001$), F2 ($t=-18.54$; $p<.001$), F3 ($t= 14.09$; $p<.001$), F1 slope ($t=3.75$; $p<.001$), F2 slope ($t= 7.76$; $p<.001$) und die Euklidischen Distanzwerte (-11.14; $p<.001$ und für Euklidische DistanzPER $t=13.49$; $p<.001$) festgestellt werden. Zu Kontrollzwecken wurde außerdem der Einfluss des Vokaltyps auf die Dauer überprüft ($t= 60$; $p<.001$). Vokale des Typs <Saat> sind signifikant länger als <satt>-Typ-Vokale. Die F1 Werte und die F3 Werte sind bei <Saat>- im Vergleich zu <satt>-Vokaltypen signifikant höher, während die Werte für F2 deutlich niedriger sind. Die Werte für F1 slope und Euklidische DistanzPER sind signifikant höher für <Saat>-Vokaltypen und signifikant niedriger für F2 slope und die Euklidische Distanz (vgl. z.B. die Darstellung der Mittelwerte mit Standardfehlern für Dauer, Euklidische DistanzPER und F2 in Abb. 1).

Mit Hilfe von logistischen Regressionsmodellen werden in der zweiten Studie die Einflüsse der Faktoren ‚F1', ‚F2', ‚F3', ‚Euklidische Distanz', ‚Euklidische DistanzPER', ‚F1 slope' sowie ‚F2 slope' und - zur Kontrolle - ‚Vokaldauer' auf die Unterscheidung des Vokaltyps überprüft und die Faktoren (hier als Prädiktorvariablen) hierarchisiert. Hierzu werden mehrere statistische Modelle berechnet und log-likelihood-Werte als Parameter zur Modellgüte verglichen. Es werden jeweils die Werte zwischen dem komplexesten Modell mit allen Prädiktorvariablen und den weniger komplexen Modellen, in denen jeweils eine Prädiktorvariable ausgelassen wird, verglichen. Je größer die Differenz zwischen dem komplexesten Modell und dem jeweils weniger komplexen Modell ist, desto mehr Varianz wird durch die weggelassene Variable erklärt.

Das komplexe Modell mit allen Variablen zeigt zunächst, dass alle Prädiktorvariablen einen signifikanten Einfluss auf die Unterscheidung des Vokaltyps aufweisen (vgl. Tab. 1).

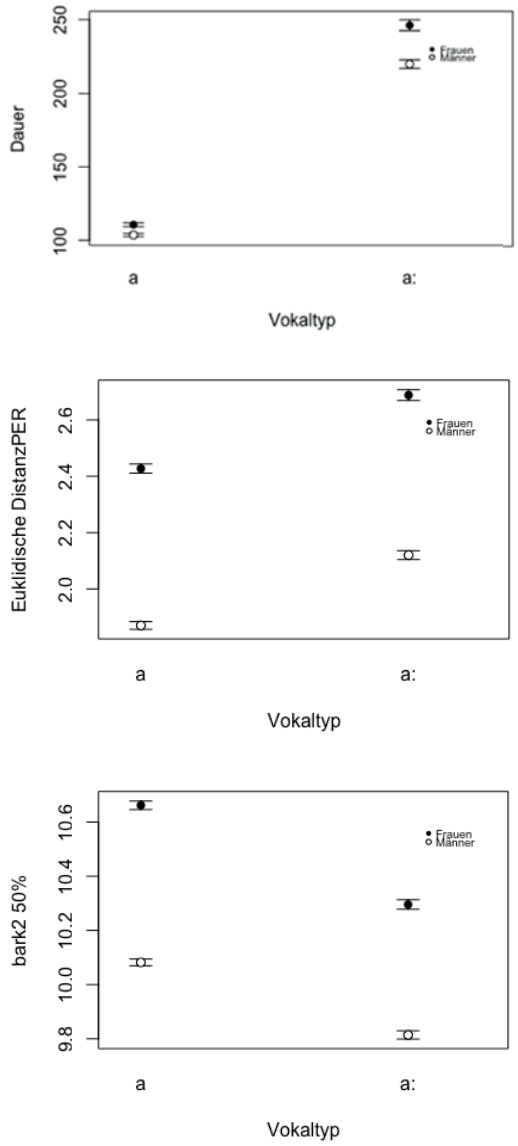

Abbildung 1. Dargestellt sind die Mittelwerte für Dauer, Euklidische DistanzPER und F2 (in Bark, bei 50%) mit Standardfehlern in Abhängigkeit der Vokaltypen und gruppiert nach Geschlecht.

Tabelle 1. Statistische Kennzahlen zum binär logistischen Regressionsmodell (Kriteriumsvariable = Vokaltyp)

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-4.2264  -0.6237  -0.3339   0.5972   3.0689

Coefficients:
               Estimate Std. Error z value Pr(>|z|)
(Intercept)    1.6190980  0.7690893   2.105 0.035273 *
b1tmid      -0.3247109  0.0900429  -3.606 0.000311 ***
b2tmid      -0.7761036  0.0662885 -11.708  < 2e-16 ***
b3tmid       0.2484469  0.0508276   4.888 1.02e-06 ***
dur          0.0225630  0.0006269  35.989  < 2e-16 ***
eucldhz2575 -0.0011610  0.0002530  -4.590 4.44e-06 ***
f1slope      0.0755948  0.0160740   4.703 2.56e-06 ***
f2slope     -0.0240989  0.0122124  -1.973 0.048460 *
eucldistcentremidbark 0.6202786 0.1034733 5.995 2.04e-
09 ***

Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’
0.1 ‘ ’ 1
```

Im Folgenden sind die Faktoren geordnet nach der Differenz zwischen dem Modell mit allen Faktoren und dem jeweiligen Modell, in dem der betreffende Prädiktor nicht mit aufgenommen wurde, aufgelistet. In Klammern stehen die Differenzen zwischen maximal komplexem Modell und Modell mit der jeweils fehlenden Prädiktorvariable:

Dauer (1283.37) > F2 (67.21) > Euklidische DistanzPER (18.88) > F3 (12.09) > F1 slope (11) > Euklidische Distanz (10.99) > F1 (6.83) > F2 slope (2.05)

D.h. neben der Dauer sind vor allem qualitative Faktoren relevant, nämlich F2 und Peripheralität.

Um regionale Einflüsse auf die Unterscheidung der Vokaltypen zu untersuchen, werden in der dritten Studie alle SprecherInnen des Deutsch-heute Korpus berücksichtigt. Hierzu werden die SprecherInnen 12 Regionen (im deutschsprachigen Raum) zugeordnet. Aus dem Korpus der Wortlisten werden die Wörter <Sprache> (=<Saat>-Typ-Vokal) und <Rache> (=<satt>-Typ-Vokal) analysiert. Ähnlich wie in der ersten Studie werden mehrere gemischte Modelle berechnet und die interindividuelle SprecherInnenvariation als Zufallsfaktor im statistischen Modell kontrolliert. Der Einfluss des Vokaltyps und der Region (als Haupteffekte und in Interaktion) wird auf die in Studie 1 beschriebenen Kriteriumsvariablen getestet (‚F1‘, ‚F2‘, ‚F3‘, ‚Euklidische Distanz‘, ‚Euklidische DistanzPER‘, ‚F1 slope‘ und ‚F2 slope‘ und ‚Vokaldauer‘). Für alle Werte außer für F3 und für die Euklidischen Distanz lassen sich signifikante Unterschiede zwischen den

Vokaltypen finden. Außerdem werden regionale Unterschiede bestätigt. Vor allem im hoch- und höchstalemannischem Raum sowie in Luxemburg unterscheiden sich die Werte in Abhängigkeit des Vokaltyps. Bei den SprecherInnen aus Luxemburg liegen die F1 Werte und die F2 Werte sowie die Euklidische DistanzPER für die beiden Vokaltypen insgesamt weiter auseinander (für F1 Vokaltyp*Region: t=.01, p<.05; für F2 Vokaltyp*Region: t=.0002, p<.005; für Euklidische DistanzPER Vokaltyp*Region: t=.03, p<.05). Die Qualitätsunterschiede sind hier im Vergleich zu den anderen Regionen größer. Im hoch- und höchstalemannischen Raum liegen die Werte für F1 und die Euklidische DistanzPER näher beieinander (für F1 Vokaltyp*Region: t=.02, p<.05; für die Euklidische DistanzPER Vokaltyp*Region: t= .008, p<.01) Die Qualitätsunterschiede sind in dieser Region geringer.

Zusammenfassung

Es lässt sich zeigen, dass neben dem phonetischen Merkmal der Quantität auch robuste qualitative Unterschiede für die Vokalopposition der deutschen <a>-Laute festzustellen sind. Des Weiteren könnten dynamische Faktoren bei der Unterscheidung eine Rolle spielen, wenngleich eher untergeordnet.

Referenzen

Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R.H.B., Singmann, H., Dai, B., Grothendieck, G. & Green, P. (2017). Package 'lme4'. R package version 1.1-13.

Boersma, P. & Weenink, D. Praat (2016). Doing phonetics by computer [Computer Programm]. Version 6.0.17, abgerufen am 21 April 2016 von http://www.praat.org/. Tübingen, 129-152.

Bohn, O., Flege, J.E., Dagenais, P.A. & Fletcher, S.G. (1992). Differenzierung und Variabilität der Zungenpositionen bei der Artikulation deutscher Vokale. *Zeitschrift für Dialektologie und Linguistik*, *72*, 1-26.

Brinckmann, C., Kleiner, S., Knöbl, R., & Berend, N. (2008). GERMAN Today: an areally extensive corpus of spoken Standard German. In N. Calzolari, K. Choukri, B. Maegaard, J. Mariani, J. Odijk, S. Piperidis & D. Tapias (Hg.), *Proc. 6th International Conference on Language Resources and Evaluation (LREC)* (S. 3185-3191). Marrakesch, Marokko: European Language Resources Association (ELRA). (http://www.lrecconf.org/proceedings/lrec2008/pdf/806_paper.pdf).

Dobson, A. J. (1990). *An Introduction to Generalized Linear Models*. London: Chapman and Hall.

Hastie, T.J. & Pregibon, D. (1992). Generalized linear models. In J. M. Chambers & T. J. Hastie (Hg.), *Statistical Models in S* (Kapitel 6). Boca Raton: Chapman & Hall/CRC.

Hoole, P. & Mooshammer, C. (2002). Articulatory analysis of the German vowel system. In P. Auer, P. Gilles & H. Spiekermann (Hg.), *Silbenschnitt und Tonakzente* (S. 129-153). Tübingen: Niemeyer.

Kohler, K. J. (1992). *Phonetisch-Akustische Datenbasis des Hochdeutschen*. Kieler Arbeiten zu den PHONDAT-Projekten 1989-1992. (=Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel (AIPUK) 26).

McCullagh, P. & Nelder, J. A. (1989). *Generalized Linear Models*. London: Chapman and Hall.

Raffelsiefen, R. (2016). Allomorphy and the question of abstractness. Evidence from German. *Morphology 26, 3*, 235-267.

Weiss, R. (1976). *The perception of vowel length and quality in German: an experimental-phonetic investigation*. (=Hamburger phonetische Beiträge 20). Hamburg: Buske

Wesker, T., Meyer, B., Wagener, K., Anemüller, J., Mertins & A., Kollmeier, B. (2005). Oldenburg Logatome Speech Corpus (OLLO) for Speech Recognition Experiments with Humans and Machines. *Proc. of Interspeech 2005*. 4-8 September 2005. Lisbon, Portugal.