

Robert Skoczek
Uniwersytet Vizja w Warszawie
ORCID: <https://orcid.org/0000-0002-1495-7116>
e-mail: r.skoczek@vizja.pl

Nuklearität der Sonorkonsonanten nach Vokalen und Diphthongen infolge der Elision des Schwa-Lautes im Standarddeutschen

**Nuklearność spółgłosek sonornych po samogłoskach i dwugłoskach
wskutek elizji głoski schwa w niemczyźnie standardowej**

**The nuclearity of sonorous consonants after vowels and diphthongs
resulting from the elision of the schwa sound in standard German**

Abstrakt

Die Synkopierung des [ə] in der Endung <-en> gilt als Indikator für die phonostilistische Differenzierung des Standarddeutschen. Selbst mit hoher Sprechspannung offiziell verlesene Texte enthalten viele Reduktionsformen von <-en>. Nach der Elision des Zentralvokals bildet der Konsonant [n] meistens den neuen Silbenkern der unbetonten Silbe. Diese Restrukturierung der Silbe kommt nicht nur auf diese Weise zustande, sondern kann zur Silbenzahlreduktion führen. Besonders umstritten ist die Nuklearität von [n] in direkter Position nach betonten Vokalen und Diphthongen. In der Fachliteratur und in den Aussprachekodizes wird die Realisation des Nasalkonsonanten uneinheitlich angegeben. Im vorliegenden Beitrag vertritt der Autor die Ansicht, dass das [n] in dieser Distribution keinen Silbenkern bilden kann. Der Verlust der Nuklearität wird aus verschiedenen Blickwinkeln begründet.

Schlüsselwörter: Nuklearität, Silbe, Sonorität, Phonostil, Elision, Sonorant, Silbischkeit

Abstrakt

Synkopa [ə] w końcówce <-en> uchodzi za indykatorem zróżnicowania fonostylistycznego niemczyzny standardowej. Nawet teksty odczytywane publicznie, cechujące się wysokim napięciem artykulacyjnym, zawierają liczne formy zredukowane tej końcówki. Wskutek elizji samogłoski centralnej spółgłoska [n] przejmuję zazwyczaj funkcję nowego ośrodka sylaby nieakcentowanej. Takie restrukturyzacje sylaby mogą zachodzić nie tylko w ten sposób, lecz także doprowadzić do zmian ilości sylab. Nuklearność [n] jest szczególnie kontrowersyjna w bezpośredniej pozycji po akcentowanych samogłoskach i dwugłoskach. W związku z tym wymowę spółgłoski nosowej ujmowano w literaturze przedmiotu oraz

kodeksach wymowy w sposób niejednorodny. W niniejszym artykule autor opowiada się za stanowiskiem, że gloska [n] w wymienionej pozycji ze względu na uwarunkowania fonetyczno-fonologiczne nie może przejąć funkcji jądra sylaby.

Słowa kluczowe: nuklearność, sylaba, sonorność, fonostyl, elizja, sonorant, syllabiczność

Abstract

The shwa sound [ə] in the ending <-en> is considered an indicator of the phonostylistic differentiation of standard German. The nuclearity of sonorous consonants after vowels and diphthongs is handled differently in the literature of the subject and in pronunciation codes. The article assumes the phonological and phonetic perspective and attempts to determine whether sonorous consonants in such a distribution are syllabic.

Keywords: nuclearity, sonority, phonostyle, elision, sonorant, syllabicity

Einleitung

Zu einem der wichtigsten Indikatoren für die phonostilistische Differenzierung des Standarddeutschen gehört die Realisation der unbetonten Finalsegmente [ən] nach Obstruenten, Sonoranten sowie Vokalen und Diphthongen. Den Ausgangspunkt der normphonetischen Erwägungen bildet die Aussage von Christian Fandrych und Maria Thurmaier in der *Grammatik im Fach Deutsch als Fremd- und Zweitsprache* (2021: 125):

In der gesprochenen Sprache wird das /e/ [sic!] in den Endungen -en, -em, -es meist in reduzierter Form ausgesprochen und kann auch ganz entfallen: *den blauη Zaun, mit blaum Kleid* (/n/ und /m/ werden silbisch realisiert); in der Endung -er wird das /e/ meist zusammen mit dem /r/ zu einem reduzierten Vokal [ə] verschmolzen.

Ungeachtet des umstrittenen phonologischen Status des im Zitat mit /e/ transkribierten Segments (vgl. Becker 2012: 38) lässt sich den von Fandrych und Thurmaier genannten Beispielen entnehmen, dass /n/ und /m/ in den Endungen <-en>, <-em> nach Vokalen und Diphthongen silbisch gesprochen werden. In den Varianten mit geminderter Aussprachespäzision werden die Nasalkonsonanten laut Meinholt (2019: 181) und dem *Deutschen Aussprachewörterbuch* (Krech et al. 2010: 110) hingegen als unsilbisch realisiert. Die Ausspracheangaben im *Duden-Aussprachwörterbuch* (Kleiner / Knöbl 2023, Teil E) geben den Lehrenden und Lernenden des Faches Deutsch als Fremd- und Zweitsprache keine eindeutige Antwort, welche normphonetischen Formen in dieser Distribution in Seminaren vermittelt werden sollen.

Da die normativen Angaben und Transkriptionen im *Deutschen Aussprachewörterbuch* (Krech et al. 2010) mit den etablierten theoretischen Ansätzen aus der Phonetik und Phonologie weitgehend übereinstimmen, möchte ich meine Positionierung auf forschungsrelevante Konzepte und

themenbezogene Modelle stützen. Im ersten Schritt wird die Tendenz zur Silbenzahlreduktion infolge der Synkopierung des Schwa-Lautes mit den stark zentrierenden Wort(gruppen)akzent und den akzentrythmischen Strukturen in Zusammenhang gebracht. Dann werden auf Basis von silbenphonologischen Prinzipien Prominenzgrade für die einzelnen Silben der Phrase *den blauen Zaun* berechnet und die reduktionsanfälligsten Segmentfolgen bestimmt. Unter Einbezug von Sonoritätsskalen wird anschließend analysiert, in welcher Distribution Sonorkonsonanten unter reduktionsfördernden Bedingungen als Silbenkerne fungieren können. Den empirischen Teil bildet eine stichprobenhafte auditiv-instrumentale Analyse der Lautsequenz *blauen* in gesprächsnahen Textzusammenhängen. Zu diesem Zweck werden einige reduzierte Formen von *blauen* aus im Internet frei zugänglichem audiovisuellem Material erhoben und phonetisch ausgewertet. Abschließend werden Ergebnisse eines Perzeptionstestes, der unter nativen und nichtnativen Probanden durchgeführt wurde, besprochen und Rückschüsse auf normphonetische Implikationen zur Transkription der Endung <-en> nach Vokalen und Diphthongen in nichtprominenten Positionen dargestellt.

1. Phonostilistik als Ausgangspunkt für die Diskussion

Die phonostilistische Vielfalt des Standarddeutschen ergibt sich aus multifaktoriellen Gegebenheiten wie sprecherbezogene Nähe ggf. Distanz, Emotionalitätsgrad und -richtung, Komplexität des Diskurses, Intention des Sprechaktes etc. (vgl. Meinholt 2019: 101). Sie wirken sich auf den Grad der phonetischen Formstrenge aus. All diese Gegebenheiten machen ein komplexes Geflecht aus, in dem lautliche Ausprägungen und außersprachliche Faktoren interdependent bleiben. Trotz der erheblichen Variabilität phonetischer Formen im Standarddeutschen ist deren Präsenz grundsätzlich vorhersehbar. Sie gelten als reguläre Ausprägung der expliziten Ausspracheformen für bestimmte situativ-kommunikative Kontexte. Unter *Phonostil* wird ein Komplex der phonetischen Mittel verstanden, die einer mündlichen Äußerung in ihrer betreffenden Form und Situation sowie in einer bestimmten Sphäre der sprachlichen Kommunikation eigen sind (vgl. Gnutzmann 1979: 47). Sie bleiben in den sprechkommunikativen Kontexten für sprachkompetente native Speaker unauffällig und sind dadurch nicht salient im Rahmen der definitorischen Merkmale einer Standardsprache.

Bewegt man sich im Sprechausdruck auf einer Achse mit den zwei Endpunkten legere Familiarität – strenge Förmlichkeit, so wurden im Hinblick

darauf in der sprechwissenschaftlichen Forschung verschiedene Schichtungen des Standarddeutschen vorgenommen. Im *Deutschen Aussprachwörterbuch* (vgl. Krech et al. 2010: 99) wird die Artikulationspräzision in drei Formstufen unterteilt:

1. sehr hohe Artikulationspräzision, die sich auf das reproduzierende Sprechen bezieht, welches bei festlichen, offiziellen Reden oft versgebunden zustande kommt,
2. hohe bis mittlere Artikulationspräzision, die ebenfalls mit dem Verlesen von vorgefertigten Texten, also manuskriptgestützt, zusammenhängt, dennoch sind die Texte internalisiert und bewegen sich oft durch ihren prosaischen Charakter in einem Konversationsstil,
3. verminderte Artikulationspräzision, die aus den wenig geplanten Sprechtexten resultiert, die bei Gesprächen, Diskussionen, Talkshows im Hörfunk und Fernsehen die Zuhörer- ggf. Zuschauernähe herstellen.

Die kategoriale Zuordnung der Sprechtexte zu dem jeweiligen Phonostil steht u. a. im Zusammenhang mit der Realisierung der unbetonten Silben <-en, -el, -em>, insbes. nach Sonoranten, Vokalen und Diphthongen. Ihre Realisationen gelten als Indikatoren für die jeweiligen Phonostile. Empirisch wurde nachgewiesen (vgl. Meinholt 1973), dass die vokallosen Formen häufiger in den niedrigeren Präzisionsstufen vorkommen. Bei den nativen Rezipienten riefen volle Formen eher Erstaunen und Inakzeptanz hervor, falls diese Silben in normalen Konversationen, in prosodisch nicht exponierten Redesequenzen mit dem Zentralvokal gesprochen würden (vgl. Hall et al 1995: 18). Die relaxiert-lässige Haltung der Sprechenden und die abnehmende soziale Distanz mit wachsender Vertrautheit fördern die Spontanität und die psychische Dynamik (vgl. Meinholt 2019: 102). Da jedes Gespräch in seiner Dynamik nicht gleichmäßig abläuft, sind auch Sprechmomente in der lautlichen Struktur vorhanden, die präzise Formen verlangen. Deswegen kann man in diesem Zusammenhang von Möglichkeiten der Reduktionismen reden.

Bezugnehmend auf das Zitat aus Fandrych und Thurmaier (2021: 125) werden somit vokallose Formen an sich nicht in Frage gestellt, sondern es fragt sich, ob das synkopierte [ə] nach Vokalen und Diphthongen zur Nuklearität der Sonoranten führt.

2. Der akzentzählende Sprachrhythmus als Auslöser von Reduktionismen

Bei der Kategorisierung der phonostilistischen Eigenschaften standarddeutscher Sprechtexte müssen suprasegmentale und segmentale Sachverhalte im engen Zusammenhang betrachtet werden. Diese starke Wechsel-

beziehung resultiert aus der zentrierenden Kraft der Artikulationsenergie der prominenten Silben auf Wort-, Syntagma- und Satzebene. Das Deutsche gehört zu den akzentzählenden Sprachen (vgl. Hirschfeld/Reinke 2018: 66; Hakkarainen 1995: 151), die mit der Isochronie-Theorie in Verbindung gesetzt werden (vgl. Richter 2008, 26). Das bedeutet, dass Rhythmusschläge in einer Redesequenz ungeachtet der Anzahl der dazwischen liegenden unbetonten Silben in ungefähr zeitgleichen Abständen in Erscheinung treten. Die Tendenz zur Isochronie ist im Deutschen folgenreich. Die gesamte Sprechenergie ruht auf den prominenzfähigen Silben, die im Verhältnis zu den benachbarten Silben viel präziser, dynamischer und lauter ausgesprochen werden. Den benachbarten Silben wird hingegen die Artikulationsenergie entzogen. Infolgedessen führen die Artikulatoren die Bewegungen träge und weniger präzise aus. Durch Akzelerationen der Redesequenzen werden redundante Momente in der Rede zusätzlich eingespart. Damit die temporären Abstände zwischen den Prominenzgipfeln in der Äußerung in etwa konstant bleiben, werden die quantitativen Unterschiede zwischen den Gipfeln durch Kürzungen, Substitutionen und Tilgungen ausgeglichen. Diese isochrone Rhythmusstruktur mit dem regulären Wechsel zwischen Stark und Schwach, Retardation und Akzeleration, hoher und niedriger Sprechspannung löst eine Zentralisierungstendenz der Vollvokale durch den Entzug der Artikulationsenergie in den unbetonten Silben (vgl. Rues 1993: 121). Allerdings bedeutet es nicht, dass die Artikulationsenergie allen unbetonten Silben gleichermaßen entzogen wird und es immer zur Restrukturierung der Silben kommt.

3. Autosegmental-metrische Begründung der Reduktionen in den Endungen [ən] und [əm]

In der Gesprächslautung können die Lautsequenzen nicht isoliert betrachtet werden, weil sie in einem melodisch-rhythmischem Kontext auftreten. Prosodische Strukturen, die den segmentalen Prozess mitsteuern, verhelfen zum Verständnis für besondere Reduktionsanfälligkeit mancher Segmente. So kommen in sachlich-neutraler Rede in *den blauen Zaun* zwei prominenzfähige Inhaltssilben <blau> und <zaun> infrage. Sie können potenziell akzentuiert werden, jedoch werden sie in der o. g. Phrase mit unterschiedlichen Prominenzgraden markiert. Die Silben <den> und <en> stehen hingegen in atonalen Positionen, die weiteren Reduktionsprozessen ausgesetzt werden können, solange die gesamte Inhaltsseite der Phrase

dadurch nicht beeinträchtigt wird. *<den>* kann sich außerdem als Klitikum einer stärken prosodischen Einheit (Silbe σ , Fuß F , Wort ω) anschließen.

Eine klitische Wortform ist eine morphologische Wortform, die sich an eine benachbarte morphologische Wortform *>anlehnt<* in dem Sinne, dass sie mit dieser ein phonologisches Wort bildet (Peters / Fuhrhop 2023: 132).

Dabei können sich Klitika zweierlei verhalten. Zum einen können sie in einer prosodisch exponierten Position in einer Äußerung stehen und Eigenschaften phonologischer Wörter vertreten, zum anderen können sie auf phrasenphonologischer Ebene mit einem anderen phonologischen Wort ein phonologisches Wort bilden, indem sie sich wie „Flexionssuffixe und vokalisch anlautende Wortbildungssuffixe“ (ebd. 133) verhalten.

Durch Pedifizierung der phonologischen Phrase *den blauen Zaun* kann das rhythmische Profil in einem terminalen Verlauf veranschaulicht und die Reduktionsanfälligkeit bzw. Reduktionsresistenz der Silben begründet werden.

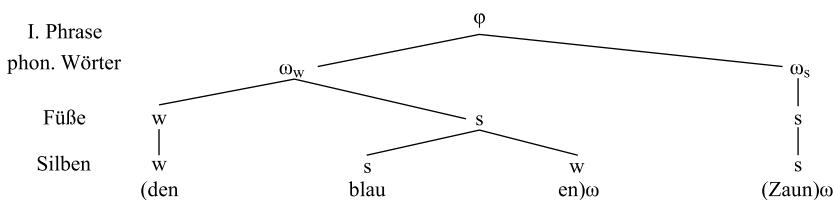


Abb.1. Baumdiagramm der phonologischen Phrase *den blauen Zaun*

In dem obigen Beispiel können die phonologischen Wörter (ω) in die übergeordnete Intonationsphrase (IP) integriert werden, zwischen denen nur bei langsamem Sprechen eine Zäsur vorkommen kann. Das bedeutet wiederum, dass mit beiden phonologischen Wörtern genauso wie auf silben- und fußphonologischer Ebene verschiedene Prominenzgrade verbunden sind. In dem Baumdiagramm werden Silben, Füße und phonologische Wörter dementsprechend mit s für *strong* und mit w für *weak* gekennzeichnet. Dem Baumdiagramm sind zwei dominante Rhythmus-Schläge auf den Vollsilben *<blau>* und *<Zaun>* zu entnehmen. Die Anzahl der direkt dominierenden S-Punkte über den einzelnen Silben bestimmt die Stellen, wo die zentrierende Artikulationsenergie kumuliert wird. Die direkten w-Punkte weisen auf den Entzug der Artikulationsenergie und die damit verbundene Abschwächung der Silben hin. Mit einem Prominenzmuster könnte die Verteilung der Artikulationsenergie unter Verwendung der Formel für dominierende S-Punkte $n + 1$ (vgl. Peters / Fuhrhop 2023: 118) nummerisch wie folgt illustriert werden:

Tabelle 1. Nominale Prominenzwerte in der phonologischen Phrase *den blauen Zaun*

Silbe	Anzahl der direkten S-Knoten	Prominenzgrad
den	0	1
blau	2	3
en	0	1
Zaun	3	4

Ersichtlich ist, dass die Silben *<en>* und *<den>* im Vergleich zu den Silben *<blau>* und *<Zaun>* mit wenig Artikulationsenergie realisiert werden und damit besonders reduktionsanfällig sind. Diese rhythmische Gestaltung der phonologischen Phrasen (ϕ) wird hier isoliert betrachtet. Hinzu wird vorausgesetzt, dass es an der Grenze zwischen ω_w und ω_s zur Retardation kommt. Bei schnellerem Sprechtempo und erhöhter Relaxation, was typisch für ein affektiertes Gespräch ist, kann die Gliederung in einer längeren Äußerung von Sprechenden variabel gestaltet werden. Eine andere Phrasierung manifestiert sich selbstverständlich in den rhythmischen Strukturen, die die Reduktionsbereitschaft und den Reduktionsgrad erhöhen oder vermindern. Hier kommt es oft zu kontextuell bedingten Prominenzverlagerungen wegen des isochronen Charakters der Rhythmisierung im Standarddeutschen. Das wird in der Äußerung *Er kauft den blauen Zaun* exemplifiziert.

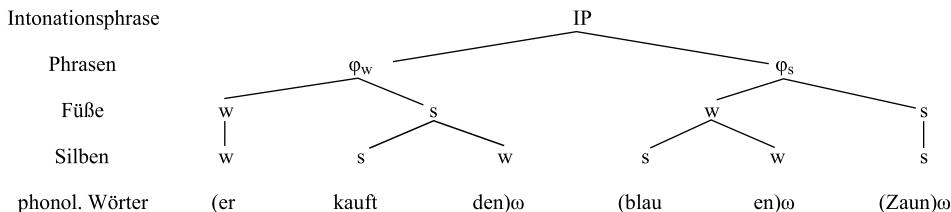


Abb. 2. Baumdiagramm der Intonationsphrase *Er kauft den blauen Zaun*

In dieser Intonationsphrase, die bei raschem Sprechtempo als eine melodisch-rhythmische Einheit realisiert wird, kann sich die Artikulationsenergie eher auf die Silben <kauf> und <Zaun> konzentrieren. Die dazwischen liegenden Silben werden zwischen diese Prominenzgipfel quasi zusammengequetscht, um den isochronen Abstand zwischen den Prominenzgipfeln zu gewährleisten. Dieser Prozess ist reduktionsfördernd. Im Gegensatz zur Silbe <blau>, die hier zum schwachen Fuß (blauen) Fw gehört, aber eine starke Silbe ist, bilden die Silben <kauf> und <Zaun> in den beiden phonologischen Phrasen starke Füße (kauft den)Fs und (Zaun)Fs und stellen in der IP reduktionsresistente Positionen dar. Die Silbe <Zaun> steht

zudem in einer starken phonologischen Phrase φs und erhält dadurch in dem Satz die höchste Prominenz, die in der Äußerung in den Fokus gesetzt wird. „Die Ausweitung des Fokus von der akzentuierten Silbe auf eine größere Einheit wird als Fokusprojektion bezeichnet“ (Peters/Fuhrhop 2023: 146). Der Fokus überragt die schwache Phrase φw , fungiert als Wortgruppenakzent (vgl. Stock 1996, 52 ff.) und gilt in der IP zugleich als Satzakzent. Am schwächsten sind in diesem Satzzusammenhang die Silben <den> und <en>. Ihnen wird die meiste Artikulationsenergie entzogen. Diese Position fördert selbstverständlich Elisionen der Vokale. Das übrig gebliebene [n] in der Endung <en> übernimmt dann eine neue Funktion in der Silbenstruktur, worauf im Weiteren eingegangen wird.

4. Nuklearität der Nasalkonsonanten in den Reduktionssilben <-en> und <-em> nach Schwa-Elisionen aus silbenphonologischer Sicht

Silben als prosodische Einheiten konstituieren, wie oben veranschaulicht, in der suprasegmentalen Hierarchie die Basis für weitere prosodische Einheiten wie Füße, phonologische Wörter, phonologische Phrasen, Intonationsphrasen oder Äußerungsphrasen. Intuitiv sind Silben allerdings im Sprachbewusstsein der Sprachteilhaber geläufig, worüber auch die silbengestützten Schriftsysteme Nachweis bieten. Kinder können relativ früh Wörter in Silben trennen.

Schwieriger als der Nachweis der weit verbreiteten Verwendung der Silbe als Analyseinheit ist das Formulieren einer allgemein akzeptierten und konsistenten Definition der Silbe (Wiese 2016: 47).

Das Problem resultiert unter anderem daraus, dass die Sprechsilbe als eine sprechsprachliche Einheit zwischen den Suprasegmentalia und Segmentalia anzusiedeln ist. Durch akustische, artikulatorische Eigenschaften der Silbenkonstituenten und deren Anordnung ist die Silbe auch als eine phonetische Größe zu betrachten (vgl. Heike 1992). Demnach lassen sich je nach dem Gesichtspunkt unterschiedliche Definitionen des Terminus *Silbe* in der Literatur finden. In Metzler-Lexikon (Glück / Rödel 2016: 716) wird sie als kleinste lautübergreifende Einheit definiert, die in phonetischer Hinsicht wahrnehmungsmäßig phänomenal gegeben ist. Zum anderen wird die Silbe als gestenbasierter Komplex begriffen, der verschiedenen Konstriktionen

im Ansatzrohr und den Ausformungen der Resonanzräume zugrunde liegt. Zur Klärung der Nuklearität bzw. Nichtnuklearität der Nasalkonsonanten nach der Schwa-Elision sind die Aspekte der segmentalen Konstruktion der Silben und der akustischen Eigenschaften im Silbengerüst von Belang. Wenn man die Silbe unter phonologischem Aspekt als eine Phonemkette betrachtet, so gelten für die Anordnung der Lautsegmente allgemeine Gesetzmäßigkeiten und Sonoritätsprinzipien.

Sonorität bedeutet hier so viel wie intrinsische Lautheit, also eine Klangfülle, die einem Segment innwohnt [...] Entscheidendes Kriterium für die Silbe ist danach der Verlauf der inhärenten Schallfülle zwischen den Segmenten, und zwar relativ zu den jeweils benachbarten Segmenten (Wiese 2016: 47).

Nach Heike (1992: 3) liegen in dem Konzept der Sonoritätshierarchie enge Bezüge zu phonetischen Phänomenen nahe, die die Abgrenzung der Silben und die Nuklearität der Segmente mitbestimmen, d. h. die Sonorität als Eigenschaft inkludiert sowohl den phonetischen als auch den phonologischen Bereich und impliziert phonetische Korrelate, die u. a. die Gestaltung von Silben mitbestimmen. Wiese (1988: 90) weist ebenfalls auf diesen Zusammenhang hin:

Die Sonorität ist einerseits eine akustisch-phonetische Eigenschaft, andererseits aber auch ein phonologisch außerordentlich bedeutsames Merkmal, da die wesentlichen Beschränkungen der Phonotaktik durch die Sonoritätshierarchie gesteuert werden.

Nach dieser intrinsischen Eigenschaft der Segmente wurden Phone in natürliche Lautklassen eingeteilt und hierarchisch geordnet. Hier beziehe ich mich auf die Sonoritätshierarchie von Bzdęga / Foss (1961: 30, 78), die in Anlehnung an Wiese (2016: 48), Becker (2012: 61) und Grassegger (2013: 70) von mir ergänzt wird.

Tabelle 2. Aufsteigende Anordnung der Sonorität in den Lautsegmenten

Sonoritätsgrad	1	2	3	4	5	6	7	8	9	10	11	12	13
Segmente	p t k ç x	s ſ f b d g z ʐ	v j z ʐ	m n ŋ	l	r R ɹ	χ ʁ ʁ χ ɻ	y u i y ɻ ɻ	ɪ ʏ ʊ ɪ ɻ ɻ	ø o e ø ɻ ɻ	ɛ ɛ ɛ ɛ ɻ ɻ	a	

Bei näherer Hinsicht ist erkennbar, dass die Sonoritätsgrade den natürlichen Lautklassen eigen sind. In der Fachliteratur werden demnach nur grobe Sonoritätsskalen angegeben, die man wie folgt darstellen kann (vgl. Becker 2013: 61; Wiese 2016: 48; Ramers 1998: 94).

Tabelle 3. Aufsteigende Anordnung der Sonorität in den Lautklassen

Sonoritätsgrad		Lautklasse	Lautunterklassen
grob	fein		
I	1	Obstruenten	Fortis-Explosive
	2		Fortis-Frikative
	3		Lenis-Explosive
	4		Lenis-Frikative
II	5	Nasale	Nasale
III	6	Liquide	Laterale
	7		Vibranten
IV	8	Approximanten	Approximanten
V	9	Vokale	hohe gespannte Vokale
	10		hohe ungespannte Vokale
	11		obermittelhohe Vokale
VI	12		untermittelhohe Vokale
	13		tiefe Vokale

Konträr zu dieser Auffassung steht allerdings die Skalierung der Sonorität bei Vater (2005: 215), der die silbischen Nasalkonsonanten als Sondergruppe auf der Skala zwischen den Nasalen und den Liquiden verortet. Dabei bleibt für mich fraglich, welche Gründe für die mittige Ansiedelung der silbenkernbildenden Segmente sprechen. Dies wird von Vater leider nicht erläutert.

Tabelle 4. Sonoritätshierarchie nach Vater (2005)

minimal	→	Sonorität			→	maximal
1	2	3	4	5	6	7
Plosive	Frikative	Nasale	silbische Nasale	Liquide	Glides	Vokale

Nach dem allgemeinen Silbenbaugesetz sollten die Silbenränder spiegelbildlich so strukturiert sein, dass die Sonorität im Onset zum Nukleus hin zunimmt und vom Nukleus weg zur Koda abnimmt (vgl. Becker 2013: 62f.). Explizite Präferenzgesetze der Silbenstruktur wurden von Vennemann (1988) formuliert. Wenn man die Skala von Vater auf das Beispiel *blauen* anwendet, entsteht folgendes Sonoritätsprofil:

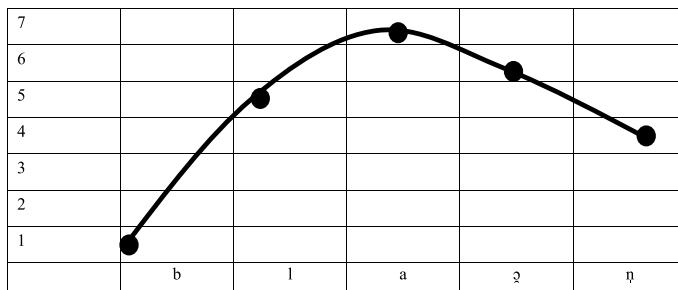


Abb. 3. Sonoritätsprofile für <blauen> in reduzierter Aussprachevariante mit silbischen Nasalkonsonanten nach der Skala von Vater (2005)

Unter Einbezug von Vaters Skala und seiner Annahme folgend, dass die Nasalkonsonanten nach der Schwa-Tilgung zu Silbenträger werden, müssten zwei Sonoritätsgipfel in Abb. 3 ersichtlich sein. Hier steigt die Klangfülle zu einem Silbenkern hin und fällt vom Silbenkern weg. Dies weist einerseits eindeutig nach, dass wir es hier mit einem Einsilber zu tun haben. Andererseits ist zu erkennen, dass die letzten drei Segmente, hier im Silbenreim, hohe Sonorität aufzeigen und die Sonoritätsschwankungen zu gering sind, um einen zweiten Silbenkern bilden zu können. Ein fragwürdiges Sonoritätsprofil ergibt sich, wenn nach dieser Skala die Sonoritätswerte den Lautsegmenten in *fallen* zugewiesen werden: 2754. Auch hier ist eher mit einer Silbe zu rechnen.

Analysiert man nach der Sonoritätsskala (vgl. Tab. 2) das Beispiel *blauen* für gehobene und geminderte Artikulationspräzision der deutschen Standardaussprache, so ergeben sich folgende Sonoritätsprofile für die beiden phonostilistischen Varianten:

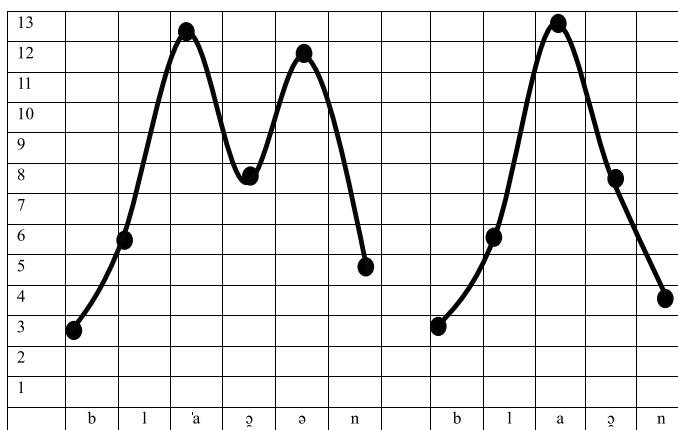


Abb. 4. Sonoritätsprofile für <blauen> in voller und reduzierter Aussprachevariante

Im ersten Fall wird die Sonorität der vollen Variante [bl'aœn] profiliert. Die vokalischen Anteile bilden hier zwei Gipfel, die gleichzeitig für zwei Silbenkerne stehen. Die Übergänge zwischen den Sonoritätsgipfeln haben eine gleichmäßig fallend-steigende Kontur. Es kommt dazwischen zu keiner weiteren sprunghaften Sonoritätsänderung. Ein solcher Verlauf der Sonorität stellt eine ideale Abbildung für die Zweisilbigkeit dar.

In der für die Gesprächslautung typischen reduzierten Form kommt nach der Elision des Schwa-Lautes eine weitgehende Silbenstrukturänderung zustande. Das [n] bildet nun nicht mehr die Koda der zweiten Silbe, sondern den Endrand der ersten Silbe. Die Sonorität verläuft steigend-fallend und im Profil ist eindeutig nur ein Höhepunkt erkennbar, weswegen der silbenauslautende Nasal laut Sonoritätstheorie keinen Nukleus in der zweiten Silbe bilden könnte, wie Fandrych und Thurmaier (2019) durch ihre Transkriptionen implizieren. Dies würde gegen die allgemeinen Silbenbaugesetze im Standarddeutschen verstößen und wäre aus phonetischer Sicht nicht sprachökonomisch. Mit einer ähnlichen Silbenzahlreduktion ist nicht nur in *blauem*, *rohen*, *auf Wiedersehen*, *schauen*, sondern auch in *fallen*, *wollen*, *sollen* zu rechnen, in denen die Sonorität der Laute zu Ende des Wortes absteigt. Eine potenzielle Einsilbigkeit tritt auch nach der Tilgung des [œ] in <-en> nach dem alveolaren Nasal auf, z. B. können [kœn]. Hierbei sind ebenfalls Längungen des Nasals vor allem vor Sprechpausen und Retardationen möglich, also [kœn:].

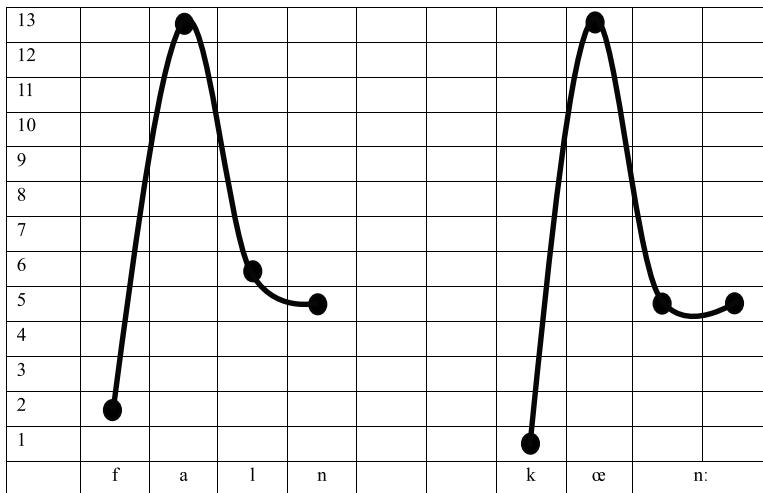


Abb. 5. Sonoritätsprofile für <fallen> und <können> in reduzierter Aussprachevariante

Ähnliche Realisationen lassen sich in der Standardaussprache mit verminderter Artikulationspräzision für *schwimmen* oder *singen* nachweisen. Nach der Elision des Murmelvokals kommt es zur progressiven Ortsassimilation des alveolaren Nasals, der durch die homorgane Bildung zunächst gelängt wird. Diese Dehnung kann natürlich bei akzeleriert-relaxiertem Sprechen reduziert werden. Dies fördert die Einsilbigkeit, welche aus silbenphonologischer Sicht auch auf die hier besprochenen Fälle in den Phrasen *den blauen Zaun* und *mit blauem Kleid* zutrifft.

5. Phonetische Korrelate bei der Silbengestaltung

Wie bereits erwähnt, ist man ohne linguistisches Wissen fähig, die Lautsequenzen zu silbifizieren und ohne Weiteres die Silbenzahl zu bestimmen. Aus strukturalistischer Sicht geschieht das auf der Parole-Ebene. Das impliziert, dass die Silbe nicht nur eine abstrakte phonologische Größe sein kann, sondern es müssen auch bestimmte phonetische Korrelate bestehen, die diese Silbifikation zu einer kinderleichten Aufgabe machen. Diesen Standpunkt vertritt auch Heike (1992: 1), der meint, dass die Silbe „auf irgendeiner empirisch verifizierbaren Ebene wissenschaftlich definiert“ sein und der phonologische Silbenbegriff mit Beobachtungen des Sprachverhaltens in Beziehung gesetzt werden kann. Deswegen stellt sich die Frage nach akusto-gestischen Vorgängen, die die Nuklearität der jeweiligen Lautsequenz konstituieren. Dabei verlässt Heike (1992: 11) das segmentdeterminierte Konzept und fasst die Silbe als ein dynamisches und ganzheitliches Phänomen auf, das alle pulmonalen, phonatorischen und artikulatorischen Gesten im Zusammenspiel ausmachen. Diese vier phonetischen Korrelate wurden schon bei der Silbenbestimmung berücksichtigt, indem die Parameter wie Expirationsstärke, Modulationen der pharyngalen Konstriktion, Kontinuität in der Luftströmung oder die Schallkraft bei der Definition als maßgebend angenommen wurden (vgl. von Essen 1979: 128–134). Keiner der Gesichtspunkte löst allerdings eindeutig das Problem der Nuklearität des [n] in *blauen* nach der Schwa-Elision.

Zur Konstanz der phonatorischen Parameter wurden vorläufige Versuche an einem Artikulationsmodell von Heike (1992) durchgeführt. Sie erwiesen einen direkten Zusammenhang zwischen Konstriktion und Schallintensität. Dennoch war der Konstriktionsgrad, der eine absolute Referenzgröße abbildet, schwer zu ermitteln. Nichtsdestotrotz konstatiert Heike, dass „im Großen durchaus ein natürlicher akusto-genetischer Zusammenhang zwischen Konstriktionsgrad und abgestrahlter Schallenergie besteht, eine

differenzierte Skalierung jedoch sehr problematisch ist“ (Heike 1992: 7). Diesen Zusammenhang sieht er in einer „durch Erfahrung erworbene[n] psychische[n] Wirklichkeit“, d. h. schon im Erstsprachenerwerb werden akzentrythmische Strukturen verinnerlicht, die den prosodischen Einheiten höheren Rangs zugrunde liegen. Da der Zugang zu diesen Strukturen erschwert bleibt, gelten laut Heike (1992: 2) die Selbstbeobachtung und Befragung der eigenen Intuition bzw. des Sprachgefühls als das effektivste Vorgehen. Alles, was Sprachteilhabende zu diskriminieren vermögen, kann erst dann auf phonetische Korrelate hin untersucht werden. Somit befindet sich der Silbenbegriff in einer Zirkularität. Laut Ramers / Vater (1995: 122) kann phonologisch gesehen mithilfe der Sonoritätsskala „die Position des Silbengipfels in Silben des Deutschen in der Regel zweifelsfrei bestimmt werden.“ Demgegenüber steht Koschmieders Konstatierung, dass diese Bestimmung der Silbengipfel aus phonetischer Sicht gar nicht so eindeutig ist, denn „wenn phonetisch eine ‚Silbe‘ vorliegt, nämlich ein Schallfüllgipfel, so braucht phonologisch dieser Gipfel nicht auch schon als Silbe zu gelten“ (1952: 290). Die Sonoritätsgipfel werden als Silben erkannt, wenn sie prosodisch relevant sind (vgl. von Essen 1979: 134), d. h. wenn ihnen eine kommunikative Funktion beigemessen wird.

Sollte bei gleicher Sonoritätskontur etwa in <blauen> [blaʊn] und <braun> [braʊn] oder <frohen> [fʁo:n] und <schon> [ʃo:n] im ersten Fall eine Silbischkeit nach der Schwa-Elision in <-en> perzipiert werden, so wirft sich die Frage auf, ob „wir hier von Nuklearität als rein psychischer bzw. mentaler Größe der Phonologie erwarten, daß sie auch [hier – RS] eine kommunikativ relevante Eigenschaft sei.“ (Heike 1992: 16). Da die Sonoritätsprofile beider Wortpaare gleich ausfallen, können diese Lautsequenzen jeweils mit einem Gipfel als einsilbig „zweifelsfrei bestimmt werden“. Heike (ebd., 17) geht davon aus, dass die Sonorität bei gleicher segmentaler Struktur nicht mehr für die Silbischkeit verantwortlich sei. Hier müssten entweder prosodische Merkmale des elidierten Nukleus oder neue Merkmale mit nuklearer Funktion im Sonorlaut auftreten. Neben dem Tonhöhenabfall und Intensitätsunterschied, die vorhin im getilgten [ən] vorhanden waren, dauert das [n] etwas länger. Das so gekennzeichnete Residuum der Vollform [bl'aʊn] kann hypothetisch den perzeptiven Eindruck von Zweisilbigkeit auslösen.

6. Vennemanns Explikation zu Silbenstrukturen im Standarddeutschen

In der Ausführung zur Silbenstruktur bezieht sich Vennemann (1982: 262) auf seine Intuition ggf. native Sprachkompetenz. Er weist auf verschiedene Silbenanzahl in den Wörtern mit gleicher Sprachlautfolge wie etwa *braun* [braʊn] vs. *blauen* [blaʊn], *Geheul* [gsh'œl] vs. *Knäuel* [kn'œl] hin. Den Unterschied sieht er in der Positionierung der Nuklearität und der Silbengrenzen (ebd., 264). Vennemann begründet seinen Standpunkt an Beispielen *Bauer*, *Feier* und *Feuer*. Der unsilbische Teil des Diphthongs [ø], [ɛ] und [œ] kommt s. E. als Silbengelenk vor (ebd., 282). Die Ambisilbizität dieser Segmente ist nachvollziehbar, denn den Diphthongen folgt ein Reduktionsvokal [ə], der als Laut mit hoher Sonorität den Nukleus der Folgesilbe bildet. Diese Wörter sind zweifelsfrei zweisilbig. Da ungespannte Vollvokale im Deutschen nie in offenen Silben auftreten, wird die silbische Komponente des Diphthongs [a] bzw. [ɔ] an das unsilbische [ø], [ɛ] und [œ] fest angeschlossen. Zur Wohlgeformtheit der Silben verhält sich der unsilbische Teil des Diphthongs wie Approximant und fungiert gleichzeitig als Onset der Folgesilbe mit dem Kern [ə], also [b'aœ̯ f'œ̯ fœ̯] analog zu [m'œ̯tə]. Dieser Meinung kann ich beipflichten, sofern dem unsilbischen Teil des Diphthongs ein sonorerer Laut, ergo ein Vokal folgt. Diese Fähigkeit zur Nuklearität wird im selben Kontext von Vennemann (282) den Liquiden und Nasalen zuerkannt: [gœ̯tœ̯ fœ̯tœ̯ knœ̯œ̯] also auch den hier analysierten Formen [blaœ̯n] und [blaœ̯m]. Diese Analogie geht m. E. zu weit, denn bezieht man sich auf die CS-Hierarchie, die Vennemann als universell zur Explikation der Silbenstrukturen bevorzugt, sind Approximanten sonorer als Liquide und Nasale. Demnach können sie in diesem Fall das Silbengelenk nicht bilden. Im Wörterverzeichnis des elektronischen Duden-Aussprachewörterbuchs (Kleiner / Knöbl 2023) wird exempli verbi das Lemma *Knäuel* als einsilbiges Wort transkribiert und gesprochen. Eine ähnliche Lautform für die Realisation des <-el> konstatiert Meinholt (vgl. 2019: 178).

7. Exemplifizierung der Realisation von *blauen*

In diesem Teil des Beitrags wird versucht, der Frage nachzugehen, ob die vorab angeführten akzentmetrischen und silbenphonologischen Gründe für die einsilbige Aussprache der ursprünglichen Zweisilber nach der Elision des Schwa-Lautes in der Endung <-en> nach Vokalen und Diphthongen in der Sprechwirklichkeit widergespiegelt werden. Dazu wurden exemplarisch

Phrasen mit den Lautformen in *blauen* mit elidiertem Schwa-Laut mithilfe dem Online-Browser youglish.com/german herausgesucht. Mit dem Programm Audacity wurden die Mitschnitte im wave-Format extrahiert und als Wavodateien registriert. Zur auditiv-akustischen Analyse wurden die Programme PRAAT und wavesurfer verwendet. Da die Silbischkeit im Lautkontinuum schwer ermittelbar ist (vgl. Kohler 1995: 76), werden die spektralen Bilder vorrangig zur Selektion der Formen mit elidiertem Schwa eingesetzt. Zur Erfassung der Silbenzahl wurde ohrenphonetisch unter Einbeziehung von Kontrollhörenden gearbeitet. Im Zweifelsfall wurden die Audiomitschnitte rückläufig abgespielt. Die rückläufige Wiedergabe soll zur Feststellung der Silbenzahl nach der Elision des Zentralvokals verhelfen.

Aus den über 40 Mitschnitten wurden nur diejenigen berücksichtigt, die folgende Kriterien erfüllt haben:

1. Die Aussprache ist ohne erkenntlich regionale Einflüsse,
2. Die Tonqualität ist analysegeeignet,
3. Tonaufnahmen sind ohne spektralbildverfälschende Musikuntermalung oder andere Geräuschkulissen,
4. Mitschnitte bilden möglichst das gesprächsnahe Sprechen ab,
5. Lautformen sind mit dem elidierten Schwa-Laut.

Insgesamt waren es 30 Audiomitschnitte. Sie stammen aus den frei zugänglichen Beiträgen entweder einst im Rundfunk ausgestrahlter Sendungen oder öffentlicher Debatten oder aus Kommentaren der Influencer auf ihren YouTube-Kanälen.

Die mit Audacity aufgenommenen Mitschnitte wurden gekürzt und enthalten nur für die hier behandelten Fälle relevante und prosodisch abgeschlossene Sequenzen. Die in den Mitschnitten befindlichen Realisationen von *blauen* wurden anschließend spektrografisch analysiert, um die vokalischen Anteile nach dem Diphthong auszuschließen. Der Lautkorpus wurde von den beiden Hörenden separat ohrenphonetisch analysiert und transkribiert. Bei Abweichungen wurde ein weiterer Experte zum Kontrollhören gebeten. Die Ergebnisse wurden zusammengefasst, wie es mit einigen Beispielen in der unteren Tabelle exemplifiziert wird.

Tabelle 5. Auszug aus Lautkorpus *blauen*

Redesequenz	IPA-Transkription	Quelle	Sender
den blauen Augen	[dən blaʊn əgŋ̩]	https://youtu.be/iHafIxsl5jw	Bairischer Rundfunk
diese blauen Lichtanteile	[dızə blaʊn l'ɪçtəntaɫə]	https://youtu.be/c-sH1O1jlxQ	Bairischer Rundfunk
aber beim Blauen	[a:βə blaʊm b'aŋ̩]	https://youtu.be/faumsGHJhrA	WDR

Das methodische Vorgehen wird an einem Mitschnitt aus einer Befragung zu einer lokalen Kunstausstellung veranschaulicht. Im Sprechverlauf einer jungen Sprecherin manifestieren sich diverse Koartikulationserscheinungen, die laut Neppert (1999: 252) weder an Silbe noch an Wort gebunden sind. Ihr Ausmaß hängt mit der Lautumgebung und mit den suprasegmentalen Faktoren wie Sprechtempo, Spannung zusammen. Durch die Korrelation von Akzeleration und Relaxation kommt es in ihrer Äußerung zur Minde rung der Sprechspannung. Das aufgrund seiner kommunikativen Redundanz von der Sprecherin Verformte bzw. Nichtrealisierte verlagert sich in der sprechsprachlichen Kommunikation auf die Antizipationsfähigkeit der Hörenden (vgl. Linder 1981: 301). Abbildung 6 stellt ein Signalkontinuum in Form eines Spektrogramms dar, in dem die rot umrahmte Sequenz die Realisierung der Abfolge /bl'aən/ illustriert. Das Sprachsignal lässt sich keine klar abgrenzbaren Laute segmentieren, da die Segmente kontinuierlich ineinander übergehen und deren Merkmale sich vor- und nachgreifend im Signal hinziehen können.

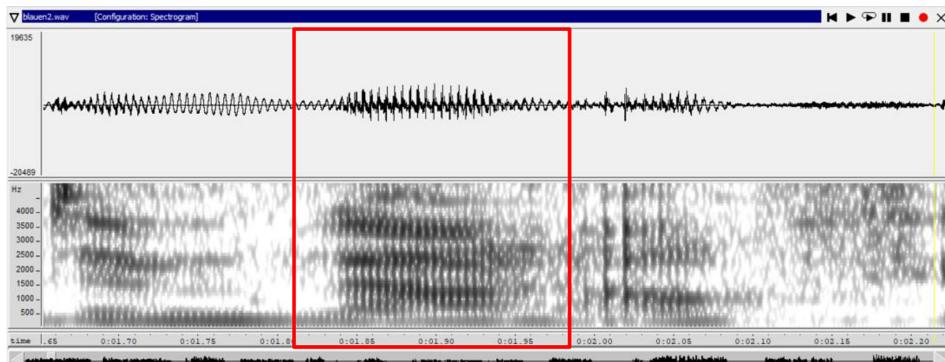


Abb. 6. Lautsequenz *<den blauen auch>* [dən bl'aən aχ]

In der Äußerung *ich find den blauen auch viel besser als jetzt grad den lilanen, weil er ein bisschen intensiver aussieht* realisiert die Sprecherin zwei Intonationsphrasen, in denen die höchste Prominenz den gegenüber gestellten Komparativformen *besser* und *intensiver* zukommt. Die zweitgrößte Prominenz in der ersten Intonationsphrase liegt auf dem Adjektiv *blau*. Die 1,974 Sekunden lange IP zählt 15 Sprechsilben. Bezugnehmend auf die Referenzdaten von Neppert (1999: 250) gilt diese Lautsequenz als sehr schnell gesprochen. In der gleichen Zeiteinheit werden Lautsegmente in einer längeren Phrase kürzer gesprochen als die in einer kürzeren Phrase. Der zeitliche Rahmen der Isochronie ändert sich je nach den genannten Bedingungen (ebd.). Um die inhaltswichtigen Silben dennoch zu exponieren und

für den Empfänger erfassbar zu machen, wurden atonale Passagen diesen prominenten Silben untergeordnet. Die einsetzenden koartikulatorischen Prozesse finden im Abgleich mit dem antizipatorischen Potenzial der Dekodierung durch den Empfänger statt. Der rückwärts abgespielte Mitschnitt wies einen Prominenzgipfel und somit die Einsilbigkeit nach. Die auditive Analyse durch weitere phonetisch geschulte Kontrollhörer bestätigte dieses Ergebnis. Probabilistisch wurde aus dem Lautkorpus ein weiterer Mitschnitt analysiert: *mit einem blauen Stift*. In dem Beispiel wurde die Form *blauen* etwas verlangsamt ausgesprochen, dennoch handelt es sich um eine Realisation ohne Schwa. Diese prosodische Struktur mit Längung des Nasals führte zu verschiedenen Hörerurteilen. Die Diskrepanzen zwischen Registrierung und Perzeption werden folgendermaßen erhellt: „Es kommen auch Wörter vor, die nur aus einem Schallgipfel bestehen, obwohl sie zweisilbig empfunden werden (z. B. fallen). Es besteht also keine ausnahmsfreie Deckung zwischen *Eindruck* und *Registrierung*.“ (Lindner 1981: 287). Für diesen Höreindruck macht Meinhold (2019: 174f.) u. a. eine prosodische Kontrastbildung durch steigenden Melodieverlauf verantwortlich. Die Zwischenform fällt dann „in die Vorbereitungsphase des [n]“. Dieser Übergang von [aɔ̯] zu [n] könnte wie folgt transkribiert werden: [bl'aɔ̯w̯n]. Die hochgestellten Lautzeichen stehen für die Überkürze der Lautsegmente und lösen den Eindruck von zwei Silben aus. Der Vokal [w̯] kann allerdings bei Retardation der Lautsegmente in der Transition als zweiter Silbengipfel fungieren. In dem Fall kann nicht die Reden von der Elision sein, sondern von der quantitativen und qualitativen Vokalreduktion. In den analysierten 30 Mitschnitten wurde der Schwa-Laut synkopiert, aber der Höreindruck unterschied sich in 12 Fällen, sodass neben der Form [bl'aɔ̯n] auch Varianten [bl'aɔ̯w̯n] gab.

Im nächsten Schritt wurde ein Perzeptionstest unter 10 nativen und 10 nichtnativen Laien durchgeführt. Ihnen wurden die Lautsequenzen [dən blaɔ̯n ts'aɔ̯n] und [dən bl'aɔ̯w̯n ts'aɔ̯n] zum mehrmaligen Abhören angeboten. Der Test ähnelt einer Identifikationsübung: *Sie hören zwei Phrasen. Sie können jede Phrase beliebig oft abspielen. Sie können das Hörbeispiel nachbrummen und den Rhythmus nachklatschen. Kreuzen Sie anschließend an, welcher Rhythmus zu der jeweiligen Phrase passt: ba ba BAM oder ba ba ba BAM. (ba steht für eine unbetonte Silbe, BAM für die betonte Silbe).*

Das Ergebnis dieser Befragung weist nur eine Tendenz auf, dass die erfragte Lautsequenz als dreisilbig gewertet wird. Dabei streute die Wertung bei den Mutersprachler:innen häufiger als bei den Nichtmutter-sprachler:innen.

Tabelle 6. Identifizierung der Silbenzahl

	[dən blaɔ̯n ts'aɔ̯n]		[dən bl'aɔ̯w̃n ts'aɔ̯n]	
	dreisilbig	viersilbig	dreisilbig	viersilbig
Muttersprachler:innen	7	3	2	8
Nichtmuttersprachler:inn	9	1	3	7

Dieser Teil der Untersuchung bestätigt nur, dass es Diskrepanzen zwischen Eindruck und Registrierung gibt. Im Nachgespräch haben viele Proband(inn)en geäußert, dass es ihnen schwerfiel, das Rhythmusmuster zu bestimmen. Auch beim Nachbrummen waren sie sich unsicher, ob sie vier oder drei Silben hören. Zum einen ist diese perzeptive Unsicherheit laut Lindner (ebd.) auf gespeicherte Grundstrukturen zurückzuführen.

Der perzeptive Eindruck führt zweifellos auch zu einem Gedächtniseindruck, der bei einer erneuten Perzeption immer wieder mit reproduziert wird und mit dem aktuellen Eindruck assoziiert wird. Dadurch ist zu erklären, daß selbst dann, wenn in unbetonten Silben der Vokal als Silbenkern objektiv fehlt, von phonetisch ungeschulten Hörern erklärt wird, sie hätten den Vokal perzipiert. Diese Analogie erklärt sich daraus, daß ausnahmslos in allen betonten und in den meisten Silben überhaupt ein Vokal als Silbenkern fungiert (Lindner 1981: 288).

Zum anderen haben die Proband(inn)en nicht immer exakt das Gehörte nachgebrummt und nachgeklatscht. Durch Reflektieren sprachen und summten sie die Sequenzen im langsameren Tempo nach und wichen vom Hörmuster ab. Dies führte dann zu einer anderen Auswertung.

Unter den nichtnativen Laien war das Ergebnis etwas klarer. Bis auf eine Person deklarierten drei Silben gehört zu haben, d.h. das Wort [blaɔ̯n] wurde in der Phrase als Einsilber perzipiert. Die Probanden, die des Deutschen nicht mächtig waren, sollten auch das Gehörte nachbrummen, nachsprechen und die Anzahl der Silben bestimmen. Da die grundlegenden „Gedächtniseindrücke“ fehlten, konnte sich der Höreindruck auf keine Referenzbasis für ihre perzeptive Analyse stützen.

Fazit

In der *Grammatik im Fach Deutsch als Fremd- und Zweitsprache* (2021: 125) weisen Christian Fandrych und Maria Thurmaier auf die Aussprache der Flexionsendungen <-en> und <-em> in atonaler Position hin. Den Transkriptionen von *blauen* und *blauem* ist zu entnehmen, dass die Beispielwörter nach dem Schwund des Schwa-Lauten weiterhin zweisilbig bleiben. Der deutsche akzentzählende Rhythmus begünstigt zweifelsfrei in dieser

Position Elisionen des Schwa-Lautes. Diese Ausspracheangaben von Fandrych und Thurmaier weichen allerdings von den silbenphonologischen und rhythmisch-metrischen Ansätzen ab. Umstritten bleibt indes die Frage, ob die reduzierte Anzahl der Vokale mit der Anzahl der Silbenkerne einhergeht oder ob die Elision nach den Vokalen und Diphthongen einen Silbenverlust herbeiführt. Bezugnehmend auf die Sonoritätsprinzipien als universelle Gesetzmäßigkeit in der Strukturierung der Silben wurde konstatiert, dass die Zweisilbigkeit von [blaɔ̯n] ggf. [blaɔ̯m] den allgemeinen Gesetzen der Wohlgeformtheit der Silben widerspricht. Nur unter bestimmten prosodischen Bedingungen kommen solche Realisationen vor. Zum anderen wurde darauf hingewiesen, dass quantitative Veränderungen des Nasals auch bei Einsilbigkeit den Eindruck von zwei Silbengipfeln auslösen können, abgesehen von Retardationen in den Übergängen von den Vokalen oder Diphthongen zu den Nasalen. Die akustisch-audiven Analysen der Hörmitschnitte weisen zweifelsfrei die Schwa-Tilgung nach. Die durchgeführten Perzeptionstests hingegen zeigen ein disparates Bild zwischen dem registrierten Hörmuster und dem Höreindruck. In diesem unentschiedenen Fall kann man sich auf etablierte normphonetische Transkriptionen beziehen. Schon in den älteren Aussprachelehrwerken im Fach DaF/DaZ, in denen gesprächsnahe Realisationen behandelt werden, werden vokallose Existenzformen mit dem Silbenverlust nach Sonoranten angegeben (vgl. Hall et al. 1995: 182; Fredrich et al. 1989: 99). In Hinblick darauf würde ich bei Prestoformen in der geminderten Artikulationspräzision folgende normphonetische Transkriptionen [d̩ blaɔ̯n tsaɔ̯n] und [mit blaɔ̯m klaɛ̯t] bewahren und sie in der *Grammatik im Fach DaF/Z* von Fandrych / Thurmaier angeben.

Literatur

- Becker T. (2012): *Einführung in die Phonetik und die Phonetologie des Deutschen*. Darmstadt.
- Bzdęga A. / Foss G. (1961): *Abriss der beschreibenden Grammatik*. Warszawa.
- Essen O. von (1979): *Allgemeine und angewandte Phonetik*. Berlin.
- Fandrych C. / Thurmaier M. (2021): *Grammatik im Fach Deutsch als Fremd- und Zweitsprache*. Berlin.
- Fredrich R.-B. / Stötzer U. / Falgowski G. (1989): *Übungen zur Phonetik der deutschen Sprache für Fortgeschrittene*. Teil I. Halle (Saale).
- Fuhrhop N. / Peters J. (2023): *Einführung in die Phonetologie und Graphematisierung*. Stuttgart / Weimar.
- Gnutzman C. (1979): *Satzphonetik und Phonostilistik*. [In:] *Phonologische Probleme des Deutschen*. Vater H. (Hrsg.). Tübingen, S. 41–62.
- Glück H. / Rödel M. (2016): *Metzler Lexikon Sprache*. 5. Aufl. Berlin.
- Grassegger H. (2013): *Phonetik. Phonologie*. Idstein.
- Hakkarainen H. (1995): *Phonetik des Deutschen*. München.

- Hall Ch. / Natunen M. / Fuchs B. / Freihof R. (1995): *Deutsche Aussprachelehre. Ein Hand- und Übungsbuch für Sprecher des Finnischen*. Helsinki.
- Heike G. (1992): *Zur Phonetik der Silbe*. [In:] *Silbenphonologie des Deutschen*. Eisenberg P. / Ramers K. / Vater H. (Hrsg.). Tübingen, S. 1–44.
- Hirschfeld U. / Reinke K. (2018): *Deutsch als Fremd- und Zweitsprache*. Berlin.
- Kohler K. (1995): *Einführung in die Phonetik des Deutschen*. 2. Aufl. Berlin.
- Krech E.-M. / Stock E. / Hirschfeld U. / Anders L. (2010): *Deutsches Aussprachewörterbuch*. Berlin / New York.
- Lindner G. (1969): *Einführung in die experimentelle Phonetik*. Berlin.
- Lindner G. (1981): *Grundlagen und Anwendung der Phonetik*. Berlin.
- Meinholt G. (2019): *Phonostilistische Ebenen*. [In:] *Sprechwissenschaftliche Phonetik. Schriften zur Sprechwissenschaft und Phonetik*. Voigt-Zimmermann S. / Lorenz F. / Hüttner A. (Hrsg.). Berlin, S. 101–264.
- Ramers K. (1998): *Einführung in die Phonologie*. München.
- Richter J. (2008): *Phonetische Reduktion im Deutschen als L2*. Baltmannsweiler.
- Rues B. (1993): *Lautung im Gespräch*. Frankfurt a. M.
- Stock E. (1996): *Deutsche Intonation*. Leipzig.
- Vater H. (2005): *Neuere Phonologie-Theorien und ihre Anwendung aufs Deutsche*. [In:] *Kwartalnik Neofilologiczny* Nr. 52. Vol. 1. Warszawa, S. 23–50.
- Vater H. / Ramers K. (1995): *Einführung in die Phonologie*. Köln.
- Vennemann T. (1982): *Zur Silbenstruktur der deutschen Standardsprache*. [In:] *Silben, Segmente, Akzente*. Vennemann T. (Hrsg.). Tübingen, S. 261–305.
- Wiese R. (1988): *Silbische und lexikalische Phonologie*. Tübingen.
- Wiese R. (2016): *Rolle der Silbe in der Lautsprache*. [In:] *Handbuch Laut, Gebärde, Buchstabe*. Domahs U. / Primus B. (Hrsg.). Berlin, S. 46–63.

