

2 VOKALVARIATIONEN

In gesprochener Sprache gibt es eine Vielzahl von Vokalrealisierungen. Vokale in gleichen Äußerungen werden von verschiedenen Personen unterschiedlich realisiert. Aber auch Wiederholungen von Äußerungen eines Sprechers weisen Vokalvariationen auf.

Im folgenden wird aufgeführt, wie wahrgenommene Vokalrealisierungen auf der symbolischen und akustischen Ebene beschrieben werden. Ziel von Beschreibungssystemen von Vokalen ist es, zwei Hauptkomponenten des Kommunikationsprozesses, die Produktion und die Perzeption, zu erfassen. In Abschnitt 2.1 wird ein in der Phonetik anerkanntes Beschreibungssystem vorgestellt, das auf den Konventionen der *International Phonetic Association* (IPA) basiert. Es werden die Zusammenhänge zwischen Artikulation und Perzeption erläutert sowie in Abschnitt 2.2 die Beziehungen zwischen Artikulation und Akustik beschrieben.

In Abschnitt 2.3 wird dann auf die Ursachen von Vokalvariationen näher eingegangen. Unterschiede in der Vokalproduktion werden im allgemeinen auf physiologische und kontextuelle Faktoren zurückgeführt. Ein großer Teil von Vokalvariationen wird unter dem Begriff Reduktion zusammengefasst. Eine Durchsicht der Literatur zeigt, dass der Begriff Reduktion von einzelnen Autoren unterschiedlich gebraucht wird und auch die Ursachen für Reduktionen kontrovers diskutiert werden.

2.1 Symbolische Beschreibungen von Vokalen

Schon seit Jahrhunderten sind Phonetiker bemüht, ein adäquates Beschreibungssystem für Vokale zu finden (für einen ausführlichen Überblick s. Ungeheuer, 1962), das die Vokalproduktion und –perzeption miteinander verbindet (Ramers, 1988). Bindeglied zwischen Artikulation und Perzeption ist die Akustik. In dem in der Phonetik gebräuchlichen Beschreibungssystem, das auf den Konventionen der *International Phonetic Association* (IPA) basiert, werden Vokale in einem Trapez dargestellt und mit quasiartikulatorischen Merkmalen beschrieben (Abbildung 1). Unterschieden werden drei Parameter:

- artikulierender Zungenteil (Vorder-, Mittel- oder Hinterzunge)
- Zungenhöhe (hoch, mittel-hoch, mittel-tief, tief)
- Lippenstellung (ungerundet, gerundet)

Die Darstellung der Vokale beruht auf dem System der Kardinalvokale (*cardinal vowels*) nach Jones (1918). Jones unterscheidet acht Kardinalvokale ([i], [e], [ɛ], [a], [ɑ], [ɔ], [o], [u]), die sich jeweils an der äußeren Seite des Vokaltrapezes befinden. Die Vokale [i] und [ɑ] wurden aufgrund

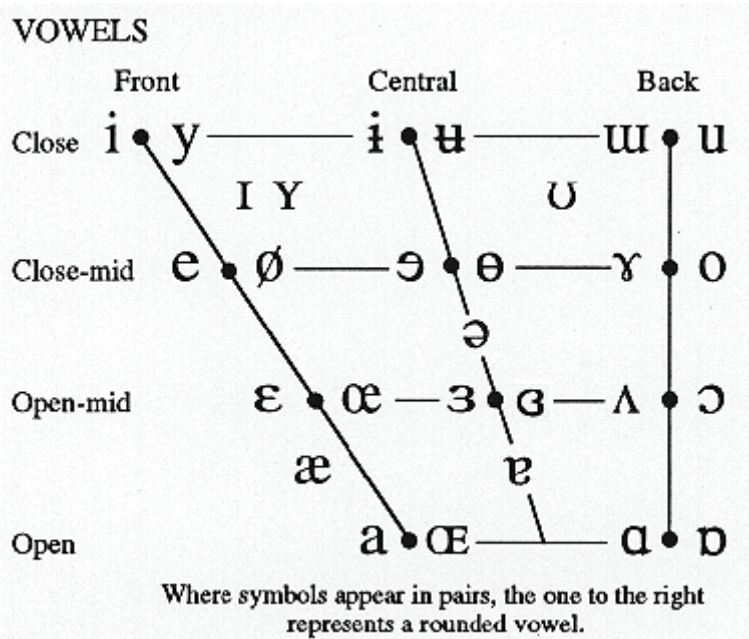


Abbildung 1: Vokaltrapez nach IPA (1993).

ihrer artikulatorischen Extremstellung als Referenzrahmen festgelegt. [i] weist als Vorderzungenvokal den geringsten und [ɑ] als Hinterzungenvokal den größten Abstand zwischen Zunge und Gaumen auf. Die anderen Kardinalvokale wurden aufgrund ihrer perzeptiven Qualität im Vokaltrapez so angeordnet, dass sie den gleichen auditiven Abstand untereinander aufweisen. Zu den primären Kardinalvokalen wurden noch sekundäre hinzugefügt ([y], [ø], [œ], [ɛ], [ɯ], [ɣ], [ʌ], [ɒ]), die sich durch die Lippenstellung – Rundung bei den Vorderzungenvokalen und dem Hinterzungenvokal [ɒ] bzw. Entrundung bei den übrigen Hinterzungenvokalen – unterscheiden.

Die Kardinalvokale sind sprachenunabhängig. Die Einteilung der primären Kardinalvokale beruht in erster Linie auf der Perzeption. Die Beziehungen zwischen Artikulation und Perzeption wurden hypothetisch festgesetzt, wobei die auditiven Abstände zwischen den Vokalen durch vergleichbare artikulatorische Abstände verursacht werden. Im Zentrum des Vokaltrapezes steht der Neutralvokal [ə], der auch als Zentralvokal bezeichnet wird. Die Mittelzungenvokale sowie [æ], [ɪ], [ɣ] und [ʊ] werden gemäß den Konventionen der IPA transkribiert, wobei letztere drei als *zentralisiert* bezeichnet werden. Dem Begriff *zentralisiert* steht der Beschreibungsparameter *peripher* gegenüber. Durch die Gegenüberstellung von *peripher* vs. *zentralisiert* wird die Abweichung der Zungenlage von der Position des Zentralvokals beschrieben. Weitere gebräuchliche Gegensatzpaare sind *geschlossen* vs. *offen* bzw. *gespannt* vs. *ungespannt*. *Geschlossen* vs. *offen* beschreibt genau genommen Abweichungen in der Zungenhöhe, vielfach werden sie aber auch mit dem Kieferöffnungsgrad assoziiert, da im Normalfall die Zungenhöhe und der Öffnungsgrad des Kiefers positiv miteinander korrelieren. Das Begriffspaar *gespannt* vs. *ungespannt* verweist auf eine unterschiedliche Muskelspannung der Artikulationsorgane (Jakobson, Fant und Halle, 1951). Der Begriff der Gespanntheit wird in Abschnitt 4.4 noch diskutiert.

Zur genaueren Beschreibung der Vokale stellt die IPA Diakritika bereit. Die Länge von Vokalen wird mit [ː] (lang), [ˑ] (halblang) oder [̘] (extra kurz) gekennzeichnet. Nuancierungen in der Vokalqualität werden durch artikulatorische Abweichungen beschrieben, die ober- bzw. unterhalb des jeweiligen Lautsymbols gesetzt werden, wie z.B. [ë] (zentralisiert), [ě] (mittel-zentralisiert), [e̝] (erhöht), [e̞] (erniedrigt), [ɔ̟] (stärker entrundet), [ɔ̠] (stärker gerundet), ([e̟]) (vorgezogene Zungenwurzel), [e̠] (zurückgezogene Zungenwurzel), [u̟] (vorverlagert), [ɪ̠] (rückverlagert), [ẽ] (nasaliert).

Das Beschreibungssystem weist jedoch zwischen Produktion und Perzeption der vorderen und hinteren Vokale eine Diskontinuität auf (Catford, 1977; Kohler, 1995b). Während die Perzeption der vorderen Vokale [i], [e] und [ɛ] auf einer linearen Abwärtsbewegung der Zunge beruht, ist für die Wahrnehmung der hinteren Vokale [u], [o] und [ɑ] die Position der Engebildung zwischen Zungenrücken und Gaumen bzw. Pharynxwand (insbesondere für [ɑ]) von Bedeutung. Da der Pharynx keine von den im Vokaltrapez verwendeten artikulatorischen Beschreibungsparametern unabhängige Dimension bildet, werden die hinteren Vokale nicht so einheitlich klassifiziert wie die vorderen. In ähnlicher Weise verhält es sich mit den aus den primären Kardinalvokalen abgeleiteten gerundeten Vokalqualitäten. Auch hier zeigt sich, dass Phonetiker in ihren Beurteilungen von gerundeten im Vergleich zu ungerundeten Vokalen stärker voneinander abweichen (Lisker, 1989; Kohler, 1995b). Hinzu kommt, dass durch verschiedene Artikulationsgesten gleiche Vokalqualitäten erzeugt werden können, was auch als artikulatorische Kompensation bezeichnet wird.

2.2 Akustische Beschreibung von Vokalen

Vokalqualitäten werden durch ihre spektralen Eigenschaften akustisch beschrieben. Von Bedeutung sind hierbei die Formanten, die als idealisierte Resonanzen des Vokaltrakts betrachtet werden. Formanten sind durch Spektralanalyse feststellbare Energieanreicherungen. Sie sind definiert durch ihre Formantfrequenz, Formantamplitude und Formantbandbreite (Heike, 1982; zur Bestimmung der Formantfrequenz s. Hess, 1972; für eine Diskussion des Begriffs des Formanten s. Maurer, 1994). Akustische Untersuchungen von Vokalen beschränken sich meist auf die Frequenzen der ersten zwei bis drei Formanten (F1, F2, F3), da die höheren Formanten als sprecherabhängig betrachtet werden und somit keinen Beitrag zur Vokalidentifikation leisten.

Die Bildung der Formanten wird mit Hilfe des Quelle-Filter-Modells (Fant, 1970) beschrieben. Formanten können auch als Resonanzen verstanden werden, die durch Umformung eines Quellen- oder Primärsignals im Vokaltrakt entstehen. Das Primärsignal wird von der schwingenden Glottis erzeugt. Die Frequenz der Schwingungen der Glottis ist identisch mit der Grundfrequenz (F0) des Primärsignals. Die zyklischen Bewegungsvorgänge der Glottis werden in der myoelastisch-aerodynamischen Theorie (van den Berg, 1958) erklärt.

Die Formantfrequenzen von Vokalen werden aus den Resonanzfrequenzen eines Vokaltrakts (Ansatzrohr) mit konstanten Querschnittsverlauf abgeleitet, wobei Ungeheuer „den Ursprung im Resonanzmechanismus des Ansatzrohres als Ganzes, genauer formuliert, im Resonanzmechanismus der ganzen vom Ansatzrohr umschlossenen Luftsäule“ (1962:80) sieht. Der konstante Querschnittsverlauf ist mit einer Konfiguration des Vokaltrakts bei der Produktion des Neutralvokals /ə/ vergleichbar. Der Neutralvokal wird als Ursprung aller Vokale gesehen, jedoch nicht als ein organischer Prozess, sondern als ein theoretisches Axiom (Ungeheuer, 1958:36). Der Neutralvokal kennzeichnet das Vokalsystem eines Sprechers und kann als Artikulationsbasis betrachtet werden.

Obwohl die Begriffe Resonanz- und Formantfrequenzen oft synonym verwendet werden, müssen die Resonanzfrequenzen des Vokaltrakts nicht immer mit den Formantfrequenzen übereinstimmen. Eine Nasalierung von Vokalen bewirkt eine zusätzliche niederfrequente Resonanz im Bereich der Grundfrequenz, häufig eine Auslöschung von F1 und eine Dämpfung von F2 und F3 (Fant, 1970; Hess, 1972). Durch artikulatorische Kompensation können verschiedene Artikulationsgesten zu dem gleichen akustischen Ergebnis führen. Die Beziehung zwischen Artikulation und Akustik ist somit nicht uneindeutig. Durch die Artikulationsstellung kann zwar auf akustische Parameter geschlossen werden, aber von der Akustik nicht auf die Artikulationsstellung. Ladefoged (1980) schränkt aber ein, dass die artikulatorische Kompensation als Ausnahmefall betrachtet werden kann, ein Rückschluss von der Akustik auf die Artikulation ist in der Regel zulässig.

Die Abhängigkeit zwischen Artikulation und Akustik zeigt sich, wenn auch stark vereinfacht, bei einer Gegenüberstellung von Vokaltrapez und der Verteilung der Vokale im spektralen Raum. Der Vergleich des Vokaltrapezes mit gegeneinander aufgetragenen F1- und F2-Werten (Formantkarte) zeigt, dass eine Änderung der Zungenhöhe eine Verschiebung von F1 bewirkt (geschlossene Vokale besitzen im Vergleich zu den geöffneten Vokalen einen niedrigeren F1-Wert). Der artikulierende Zungenteil verursacht eine Änderung in der Lage des zweiten Formanten (Vorderzungenvokale besitzen einen höheren F2-Wert als Hinterzungenvokale). Eine Änderung der Lippenstellung wirkt sich diagonal zur F1- und F2-Achse aus.

Vokale besitzen inhärente akustische Eigenschaften, deren Ursprung in den Artikulationsbewegungen gesehen werden. Hierzu gehören die intrinsische Grundfrequenz und die Dauer. Geschlossene Vokale besitzen eine höhere Grundfrequenz als offene Vokale. Die Ursache für dieses Phänomen ist noch nicht eindeutig nachgewiesen (Whalen et al., 1998). Ohala (1973) vermutet, dass die Zugkraft der Zungenmuskeln bei höherliegender Zungenstellung auf die vertikale Spannung der Stimmlippen wirkt (*tongue pull theory*). Manchmal wird die intrinsische Grundfrequenz auch auf eine höhere Luftströmung bei geschlossenen als bei offenen Vokalen zurückgeführt (Neppert und Péturson, 1986). Mit einer erhöhten Luftströmung ist eine stärkere Strömung in der Glottis verbunden, so dass geschlossene Laute höhere F0-Werte aufweisen.