

**ZUM AFFRIKATENSTATUS DER SUPRASILBISCHEN  
SEGMENTVERBINDUNG „□ǰ“ IM DEUTSCHEN. PROBLEME UND  
EIN AKUSTISCHER LÖSUNGSANSATZ  
0 GEGENSTAND, RAISON D'ÊTRE DES THEMAS,  
HYPOTHESENSTRUKTUR UND ZIELSETZUNG DER  
UNTERSUCHUNG**

Siti Kudriyah  
Balázs Huszka  
Fakultas Bahasa dan Seni  
Universitas Negeri Medan  
Sekolah Pascasarjana Program Studi Linguistik  
Universitas Sumatera Utara

**ABSTRAK**

Die Affrikaten gehören gewiss zu den meistdiskutierten Themen in Phonetik und Phonologie; die Beschäftigung mit ihrem Wesen und Status innerhalb des Phon(em)systems einer Sprache wird von Zeit zu Zeit *en vogue* in der Linguistik.<sup>1</sup> BALASSA (1904: 129 f.) hält sie, nebst den Aspiraten (!), für „mitlautende Diphthonge“, und lässt sich über sie folgendermaßen aus:

Der Übergang vom Explosiv zum folgenden Selbstlaut kann sich auch so ereignen, dass der Verschluss in der Mundhöhle [...] nicht ganz öffnet, sondern nur so, dass die Luft sie als eine schmale Öffnung durchläuft. [...] Einen solchen Diphthong, der aus einem Explosiv wie einer ihm im Hinblick auf den Bildungsort entsprechenden Spirans besteht, nennen wir Affrikata. [...] [S]o eine Mitlautverbindung, woselbst die zwei Laute zu zwei [...] Silben gehören [...], [können wir für nicht für eine Affrikata halten] (unsere Übersetzung [aus dem Ungarischen], B.H./S.K.).

Im Zitat liegen alle drei konstitutiven Merkmale (*M*) der Affrikaten zutage, d.h., dass (i) sie sich aus einem Explosiv und einem unmittelbar darauffolgenden homorganen (= am [annähernd] gleichen Ort gebildeten) Frikativ (= Spirans; in BALASSAS Terminologie) zusammensetzen, (ii) zwischen denen sich keine holotische Verschlusslösung ereignet (bzw. der Verschluss in den Frikativ gelöst wird), (iii) resp. dass ihre zwei Segmente jeweils in eine Silbe fallen. Zur Abgrenzung der Affrikaten von den einfachen Segmentverbindungen *Explosiv + Frikativ* dient *M* (iii), das zugleich *M* (ii) als artikulatorische Notwendigkeit impliziert (und umgekehrt): Die Abbrechung expiratorischer (= Ausatemungs-)Kontinuität evtl. das damit verbundene Decrescendo (= Verminderung des Druckes, des „Expirationsstoßes“ *peu à peu*) sind Anzeichen eines Silbenauslautes, während das ihm folgende *ex abrupto* eingetretene Crescendo einen Silbeninlaut zu verstehen gibt (vgl. etwa BÁRCZI 1957: 36). Dementsprechend ist die Segmentverbindung ◉□ǰ❁ im Wort ◉□◉ǰ◆◉❁❁ eine Affrikata (sie verfügt über alle drei *M*-e), im Wort ◉ǰ&◉◉□◉ǰ❁❁ keine Affrikata (sie verfügt lediglich über *M* (i)). Leider wird in gängigen Phonetiken und Aussprachelexika nicht genug betont, dass die zwei Segmente einer Affrikata zu einer Silbe gehören; man trifft häufig auf Formen wie *dt.* \*◉ǰ&◉◉□◉◉❁❁ (DA 468) (richtig: ◉ǰ&◉◉◉□◉◉ǰ❁❁) oder *ung.*

<sup>1</sup> s. z.B. den von HORGER 1935 angesetzten „ungarischen Affrikatendisput“ über die monophonische Wertung; vgl. dazu: HORGER, Antal (1935): Mi az affrikáta? In: Magyar Nyelv 31. S. 210–218

\* $\ominus \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow$  („Schneebeesen“) (BOLLA 1995: 44) (r.:  $\ominus \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow \updownarrow$ ). Sonach, unsere erste, nomologische (= gesetzartige, vgl. dazu HÄRSING 1981: 66) Hypothese (H1) ist es, dass die Segmentverbindung von dem Explosiv  $\ominus \square \ominus$  und dem unmittelbar darauf folgenden Frikativ  $\ominus \updownarrow \ominus$  in eine Silbe (d.h. in einen „Expirationsstoß“) fallend (nunmehr: die Affrikata  $\ominus \square \ominus \updownarrow \ominus$ ) anders, (H1<sub>a</sub>) physiologisch mit inkompletter Verschlusslösung (VL), (H1<sub>b</sub>) akustisch mit geringerer Amplitude des *Burst* (= VL-Geräusches) erzeugt wird als diejenige, bei der die obigen Segmente in zwei Silben fallen. Da das Primate bei der Erzeugung des Sprechschalls unangefochten die Physiologie (~ Artikulation) besitzt, und das Akustikum lediglich deren Ausfluss, hörbare Projektion ist, impliziert (H1<sub>a</sub>) als Antezedens (A) (H1<sub>b</sub>) als Sukzedens (S), d.h. (H1<sub>a</sub>) → (H1<sub>b</sub>). Von diesem Ergebnischarakter von (H1<sub>b</sub>) her werden wir mittels akustischer Schallanalyse Schlüsse auf die Bildungsphysiologie ziehen (können), ohne sie einer zusätzlichen Untersuchung zu unterziehen. Es sei indes unbedingt festgehalten, dass, trotz der theoretischen Möglichkeit der Reversierbarkeit der Implikation (bes. vonseiten des Rezipienten)<sup>2</sup>, auch andere Antezedenzen (stimmopathologische Phänomene, z.B. expiratorische Dyspnoe [= erschwerte Ausatmung], und sprachökonomische Prozesse, v.a. die Lenition<sup>3</sup>) (H1<sub>b</sub>) zugrunde liegen mögen; diesen Faktoren aber wird nicht Rechnung getragen – sie können durch auditive Kontrolle aus dem Sprechmaterial (Korpus) ausgefiltert werden. Unsere zweite (partiell durch (H1) implizierte) Hypothese (H2) ist es, dass die Bildungsdauer (BD) der Affrikata  $\ominus \square \ominus \updownarrow \ominus$  kürzer ist als die der entsprechenden Segmentverbindung  $\ominus \square \updownarrow \ominus$ , d.h.  $BD - \ominus \square \ominus \updownarrow \ominus < BD - \ominus \square \updownarrow \ominus$  (auch wenn nicht in einem 1:2-Verhältnis). (H2) ebenfalls wird mittels akustischer Schallanalyse verifiziert.

### **KLÄRUNG EINES WEITEREN ZENTRALEN BEGRIFFS: DIE SILBE**

Die Sprechereinheit, die sich innerhalb von einer Expirationsgruppe, in NEPPERTS Formulierung (1999: 161) innerhalb von „(...) eine[r] nicht durch Atmungsvorgänge unterbrochene[n] Äußerung (...)“ (= „Ausatmungseinheit“) (i) physiologisch als (eine relative, d.h. eine im Vergleich mit der Lautumgebung nachweisbare) temporäre subglottale Druckerhöhung und (ii) akustisch als (ii/a) (eine relative) Erhöhung von Schallpegel und Periodenfrequenz und (ii/b) (eine relative) Verlängerung des vokalischen Kerns (d.h. des betonten Vokals) auszeichnet, ist *eine betonte Silbe*. Zweifelsohne manifestieren sich (i), (ii/a) und (ii/b) auch auf der auditiven Ebene, was eine intuitive Silbifizierung des Wortes (d.h. Segmentierung in Silben) ermöglicht. Das Wort *suprasilbisch* im Titel des vorliegenden Beitrags bedeutet „in eine größere Einheit als die Silbe fallend“, aktuell: „in zwei aufeinander folgende Silben fallend“.

### **KORPUS UND METHODIK DER UNTERSUCHUNG**

<sup>2</sup> Die Vertreter der Motor-Theorie der Sprechperzeption von Liberman u.a. (Haskins-Forschern) postulieren, dass der Perzipient das Gehörte unbewusst auf eigene artikulationsmotorische Gesten (d.h. hirngespeicherte organokinematische Referenzmuster) zurückführt, in Form einer stummen Mitartikulation (vgl. BRENNER/HUSZKA/WERK-MARINKÁS 2006: 149).

<sup>3</sup> SZENDE (1997: 167) definiert die ~ folgendermaßen: „[sie] ist [...] die an einen polaren Gegenpunkt haltende Tendenz der Realisation, die zum Preis des Verlustes [in, B.H.] der Auseinanderhaltung der sprachlichen Zeichen die auf ein Element projizierten Auslenkungen der artikulatorischen Bewegungen reduziert (unsere Übersetzung [aus dem Ungarischen], B.H./S.K.)“.

Bei der Zusammenstellung des Korpus wurden die folgenden Prinzipien berücksichtigt: (i) Die VP-en sollten eine normgerechte, d.h. dem GWddA<sup>4</sup> möglichst nahestehende Aussprache haben; GWddA, weil (i/a) dieses die deutsche Sprechwirklichkeit (größtenteils) immer noch widerspiegelt (auch wenn es die Tendenzen der letzten 30 Jahre nicht mehr beinhaltet) und weil (i/b) es überregional (von diatopischen [= territorialen] und diastratischen [= den einzelnen Gesellschaftsschichten entsprechenden] Hypertrophien einigermaßen frei) ist. (ii) Die VP-en sollten eine akustisch gut analysierbare Stimme haben, d.h. (ii/a) keine extrem hohe  $F_0$  (= Schwingungsfrequenz der Stimmlippen) oder (ii/b) keine Dysphoniae (= Stimmbildungsstörungen) haben. (iii) Wegen der Transparenz der Untersuchung bzw. der dazu erforderlichen Veröffentlichung von dynamischen akustischen Registraten ist die Darstellung der Daten einer einzigen VP möglich; für ihre Auswahl galt, dass ihre Aussprache (nach vorherigen Untersuchungen) fähig war, Tendenzen innerhalb von einer größeren Menge zu repräsentieren.

Die technischen Daten sind wie folgt: \*.wav, 16 Bit/44100 Hz/mono. Für die akustische Analyse wurde die Software *Praat. Doing Phonetics by Computer*, Version 5.1.23 herangezogen (für den Link der Programmierseite s. Literaturverzeichnis). Die zu verbalisierenden Analysedaten lieferten Oszillogramme (*O*), Schallpegelverlaufskurven (*SchPVK*), Zeitverlaufspektrogramme (*ZVS*) und stationäre Spektren (*sS*). Das (*O*) „(...) stellt direkt die physikalische Schallwirklichkeit dar (...) [und e]s gibt den tatsächlichen *Schwingungsvorgang* der Sendemembran, der Luftteilchen (...) und der Empfangsmembran wieder“ (NEPPERT 1999: 98; Hervorhebung im Original, B.H./S.K.). Die Ordinate der Funktion von (*O*) bildet die positive/negative Spitzenwerte der Auslenkungsamplitude des schwingenden Mediums (s.o.) ab, die Maßeinheit ist Pascal (*Pa*); die Abszisse ist die Zeitachse der Schallbildung, die Maßeinheit ist Millisekunde (*ms*). Mit Hilfe von (*O*) kann man die Dauerrelationen und den Schallartenwechsel (u.a.) zwischen benachbarten Segmenten ermitteln (vgl. ebd.). Die Ordinate der Funktion von (*SchPVK*) bildet den Schall(intensitäts)pegel (*SIP*) ab, die Maßeinheit ist Dezibel (*dB*); die Abszisse ist die Zeitachse der Schallbildung, die Maßeinheit ist Millisekunde (*ms*). Der (*SIP*) ist „(...) diejenige Energiemenge, die pro Sekunde auf einer senkrecht zur Ausbreitungsrichtung befindlichen Fläche von 1 m<sup>2</sup> ankommt“ (ebd. 78); der Nullpunkt (= „Nulldezibel“) unserer Skala liegt bei 10<sup>-12</sup> W/m<sup>2</sup>, wo ein jugendlicher Normalhörer eine Sinuswelle von 1000 Hz – nach psychoakustischen Untersuchungen – zu hören beginnt. Mit Hilfe von (*SchPVK*) kann man die Intensitätsverhältnisse (u.a.) zwischen benachbarten Segmenten ermitteln. Das (*ZVS*) ist eine „quasi-räumliche“ Schalldarstellung: Die Ordinate bildet die Frequenz (*F*), die Schwingungszahl pro Sekunde ab, die Maßeinheit ist Hertz (*Hz*); die Abszisse ist die Zeitachse der Schallbildung, die Maßeinheit ist Millisekunde (*ms*); die Applikate (als Schwärzungsgrad in den Registraten) stellt den (*SIP*) dar, die Maßeinheit ist Dezibel (*dB*). Mit Hilfe von (*ZVS*) kann man die zu den einzelnen Frequenzbereichen gehörenden Intensitätsverhältnisse in der Zeitfolge ermitteln. Die (waagrecht liegenden) schwarzen Balken/Areale („Formanten“) tragen die Information über den Zusammenhang von (*F*) und (*SIP*), auf deren Grundlage die Dekodierung als Sprachelement geschieht. Das (*sS*) ist ein Ausschnitt (engl. *cross section*) aus (*ZVS*); die Ordinate bildet (*SIP*) ab, die Maßeinheit ist Dezibel (*dB*); die Abszisse gibt die Frequenz an, die Maßeinheit ist Hertz (*Hz*). (Die Zeitachse fällt hier natürlich weg!) Mit Hilfe von (*sS*) kann man die zu den einzelnen Frequenzbereichen gehörende genaue Intensität ermitteln. Um eine klare Unterscheidung zwischen der Affrikata ☉☐☉☞☞

<sup>4</sup> KRECH, Eva-Maria/KURKA, Eduard/STELZIG, Helmut et al. (Hrsg.) (1982): Großes Wörterbuch der deutschen Aussprache. Leipzig.

und dem Cluster („Pseudo-Affrikata“)  $\ominus \boxtimes \overset{\curvearrowright}{\downarrow} \star$  zu treffen, werden wir die akustische Struktur nicht nur (i) der einzelnen Segmente, sondern auch (ii) die der Transition (= des Übergangs) analysieren; für die genaue Begründung (bzw. den grundlegenden Unterschied) s. P. 0.

## DURCHFÜHRUNG DER UNTERSUCHUNGEN

Die zwei Beispielwörter sind (i)  $\ominus \boxtimes \diamond \blacktriangleleft \square \blacksquare \star$  und (ii)  $\ominus \boxtimes \diamond \blacktriangleleft \square \blacksquare \Pi \star$ ; unserer Ansicht nach ist der Cluster (C) [p] + [f] in (i) eine Affrikata und in (ii) lediglich eine „Pseudo-Affrikata“. Die (mit Praat 5.1.23 angefertigten) Registrate können im Anhang eingesehen werden. Die Bildungsdauer (BD) des ersten (C) beträgt 100.92 ms, die des zweiten 230.47 ms. Die lineare Intensitätsstruktur (IIS) des ersten (C) ist wie folgt: Sie beginnt bei 66.54 dB und steigt bzw. erreicht ihr Maximum von 68.98 dB bei 13 ms, dann fällt sie fast allmählich bis zu 65.87 dB, ihrem Minimum bei 74.44 ms ab. Die Tendenz ändert sich bei 84.90 ms – von hier steigt die (IIS) allmählich bis auf 78.46 dB. (Das ist wegen der akustischen Struktur des nachfolgenden Diphtongs so.) Auf dem (O) ist keine „richtige“ Verschlusslösung erkennbar; diese geht in ein „Frikationsrsauschen“ über (s. M (ii) in P. 0). Die (IIS) des zweiten (C) ist wie folgt: Sie beginnt bei 68.93 dB und fällt allmählich bis zu 54.73 dB, ihrem ersten Minimum bei 30.79 ms ab; dann steigt sie bis zu 57.33 dB bei 46.82 ms. Das nächste Minimum von 53.53 dB erreicht sie nach einem Abfall bei 73.45 ms. Das nächste Maximum von 63.73 dB erreicht sie bei 116.24 ms. Dann fällt sie wieder – auch wenn nicht mehr ganz allmählich – ab; bei 208.07 ms/59.90 dB fängt sie wieder an zu steigen, bis auf 65.59 dB. (In dieser letzten Periode sind auch Spuren von Stimmlippenschwingungen zu beobachten.) Auf dem (O) ist die Verschlusslösung gut sichtbar (s. M (ii) in P. 0). Das (ZVS) und die (sS) des ersten (C) liefern die folgenden Daten: Der Frequenzbereich mit der größten Intensität liegt zwischen 9100 und 12100 Hz mit 24.1/32.3 dB; doch im letzten Drittel der (BD) sinkt der Frequenzbereich mit der größten Intensität bis zu 1550 Hz und die Intensität liegt bei 13.1 dB. (Das ist wegen der akustischen Struktur des nachfolgenden Diphtongs so.) Das (ZVS) und die (sS) des zweiten (C) liefern die folgenden Daten: In der ersten Hälfte der (BD) liegt der Frequenzbereich mit der größten Intensität ganz unten auf der Frequenzskala, um 200 Hz – dank der in der stummen Haltephase des Verschlusses von Zeit zu Zeit erscheinenden subglottalen Druckschwankungen (und der damit verbundenen Stimmlippenschwingungen); die Intensität liegt zwischen 13.2 und 33.5 dB. In der zweiten Hälfte der (BD) liegt der Frequenzbereich mit der größten Intensität zwischen 8266 und 9820 Hz mit 14.3/26.2 dB.

## SCHLUSSFOLGERUNGEN

Anhand der im Anhang angeführten Registrate und der akustischen Daten unter P. 3 lässt sich Folgendes sagen:

**(HI)**, d.h., dass der Cluster von dem Explosiv [p] und dem unmittelbar darauf folgenden Frikativ [f] in eine Silbe fallend anders, ( $HI_a$ ) physiologisch mit inkompletter

Verschlusslösung (*VL*), (*HI<sub>b</sub>*) akustisch mit geringerer Amplitude des Burst (= *VL*-Geräusches) erzeugt wird als derjenige, bei dem die obigen Segmente in zwei Silbe fallen, **ist verifiziert worden**: Bei ☹☐☉↗❄ ist keine „richtige“ (*VL*) auf dem Oszillogramm erkennbar; diese geht in ein „Frikationsrauschen“ mit geringer Amplitude über, was eine inkomplette (*VL*) supponiert.

(*H2*), d.h., dass die Bildungsdauer (*BD*) der Affrikata ☹☐☉↗❄ kürzer ist als die des entsprechenden Clusters [pf], **ist verifiziert worden**: Die (*BD*) von ☹☐☉↗❄ beträgt 100.92 ms und die (*BD*) von [pf] beträgt 230.47 ms. (Dieser Unterschied ist so groß, dass die Position im Lautkörper [d.h. am Anfang/in der Mitte/am Ende] darauf keinen Einfluss hat.)



## LITERATURVERZEICHNIS

BALASSA, József (1904): Magyar fonétika. A hangok és a beszéd fiziologiai elemzése. 2., völlig umgearbeitete Auflage von „A Phonetika Elemei“. Budapest. (Magyar Hangtan; 1).<sup>5</sup>

BÁRCZI, Géza (1957): Fonetika. 2. Auflage. Budapest. (Egyetemi Magyar Nyelvészeti Füzetek).

BOLLA, Kálmán (1995): Magyar fonetikai atlasz. A szegmentális hangszerkezet elemei. Budapest.

BRENNER, Koloman/HUSZKA, Balázs/WERK-MARINKÁS, Csaba (2006): Deutsche Phonetik. Eine Einführung. Budapest.

DA = MANGOLD, Max in Zusammenarbeit mit Dudenredaktion (Bearb.) (2000): Duden Aussprachewörterbuch. 4., neu bearbeitete und aktualisierte Auflage. Mannheim/Leipzig/Wien etc. (Duden; 6).

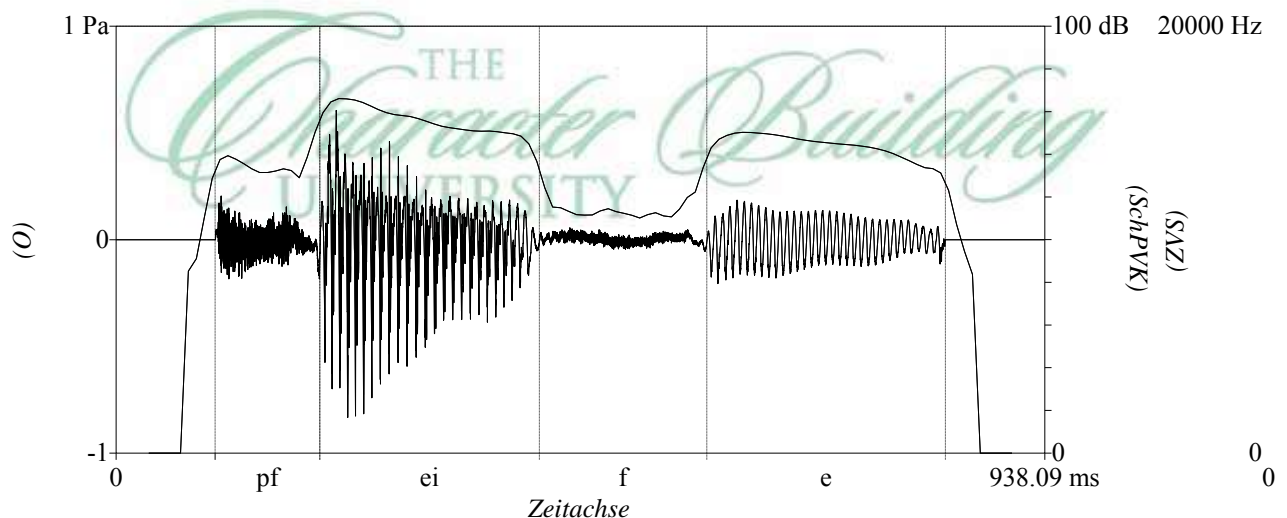
HÁRSING, László (1981): A tudományos érvelés logikája. Budapest.

NEPERT, Joachim M.H. (1999): Elemente einer Akustischen Phonetik. 4., vollständig neu bearbeitete Auflage. Hamburg.

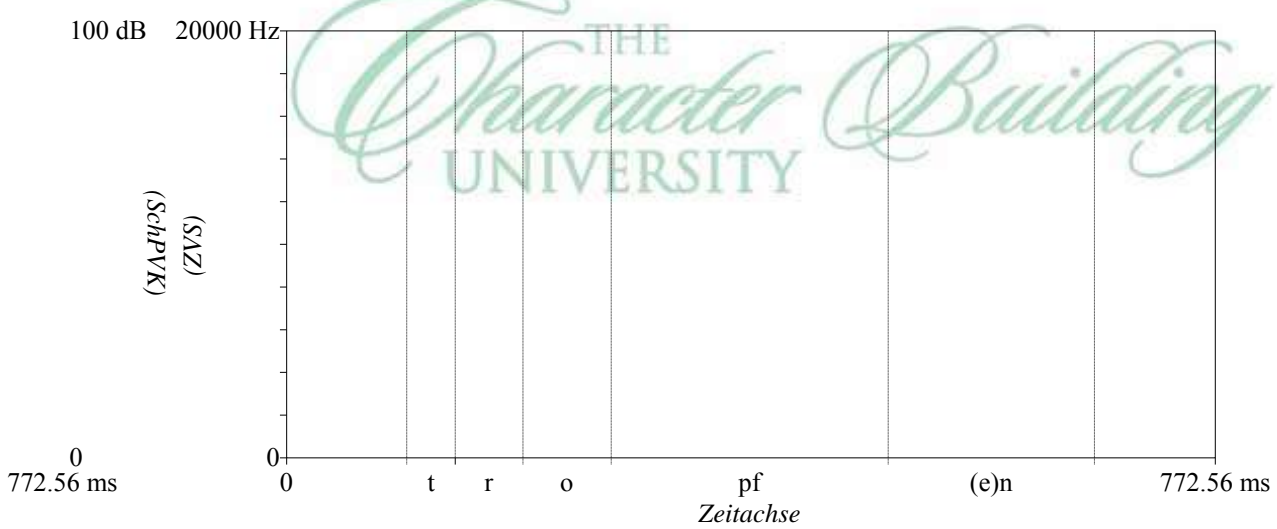
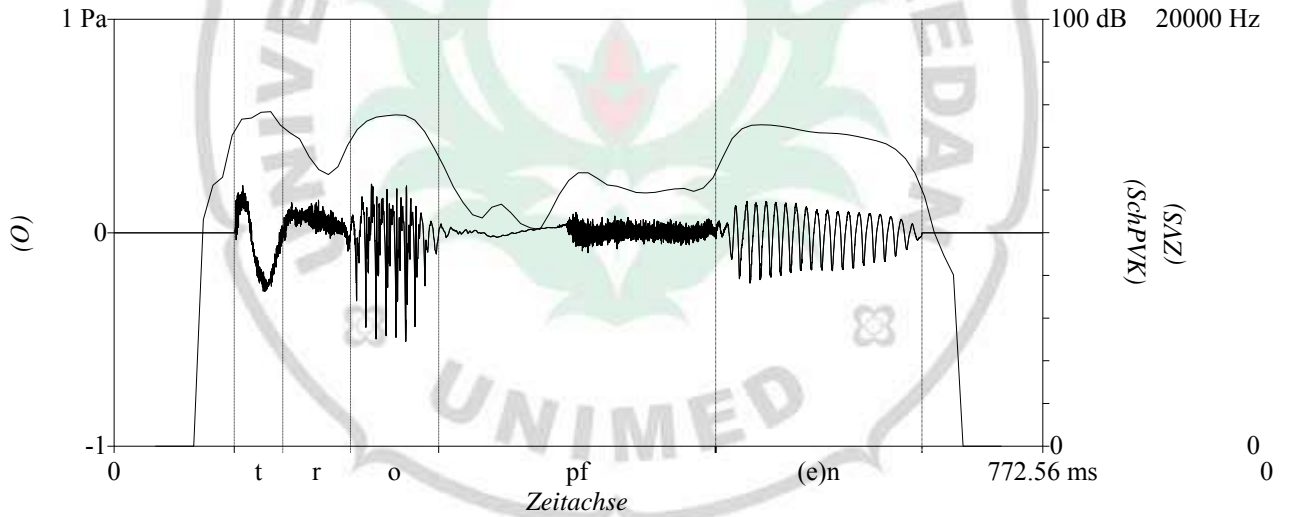
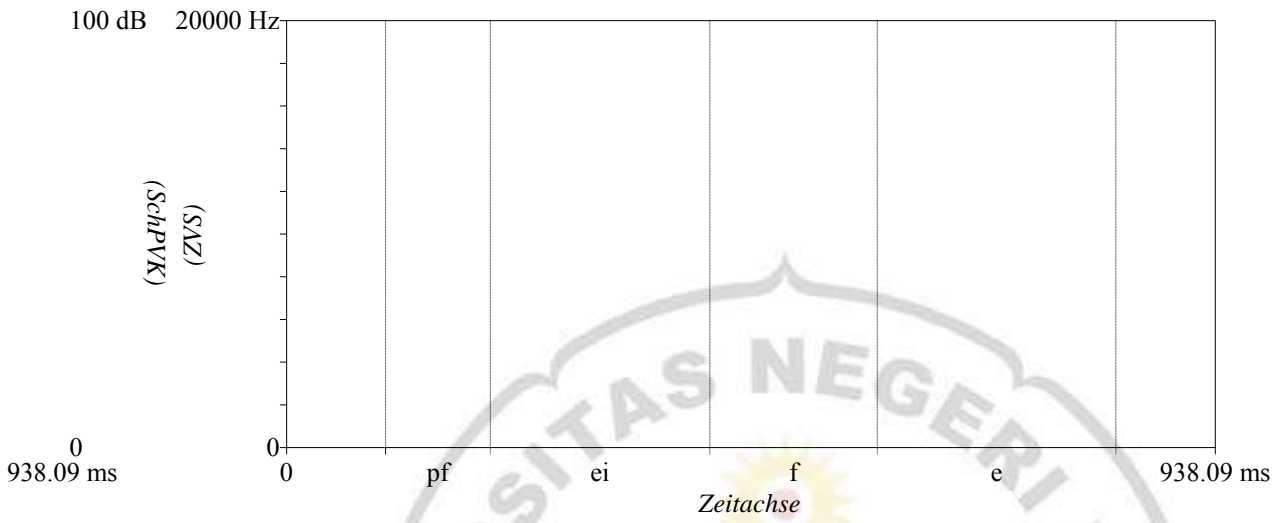
## Software

BOERSMA, Paul/WEENINK, David (2009): Praat. Doing Phonetics by Computer, Version 5.1.23. [http://www.fon.hum.uva.nl/praat/download\\_win.html](http://www.fon.hum.uva.nl/praat/download_win.html) (ges. am 13.01.2010).

## Anhang



<sup>5</sup> Bei älteren ungarischen Schreibweisen setzen wir die Bemerkung [*sic!*] nicht hinzu.





THE  
*Character Building*  
UNIVERSITY