

Bistra ANDREEVA

(Uniwersytet Kraju Saary)

Silvia BONACCHI

(Uniwersytet Warszawski)

Freundlich oder feindlich? Zur illokutionären Struktur und phonetischen Realisierung von indirekten supportiven und derogativen Sprechakten am Beispiel von Scheinbeleidigungen (*mock impoliteness*)

Abstract:

Friendly or hostile? A pilot study about the illocutionary structure and phonetic realization of indirect supportive and derogative speech acts on the example of banter expressions (*mock impoliteness*)

The “Banter Principle” describes cases in which an offensive utterance (for example: DT: “Du Arsch!” or PL: “Ty draniu”) is not addressed by the speaker to the interlocutor with an offensive intention, but it is intended to be an expression of admiration which reinforces the relationship with the Addressee. In addition, use of such language reinforces social ties, i.e. identity and a sense of affiliation to the group. The appropriate reconstruction of the intended meaning (derogatory or supportive meaning, i.e. genuine or mock impoliteness) depends on the conversational setting, on the relation of the speakers, on the mental presuppositions of the interlocutors, and on the mutual acceptance of the communicative means. The use of banter utterances is quite unstable, since they can always switch from a face-enhancing to a face-threatening or aggressive act. The paper presents the results of a pilot studies on phonetic cues of banter utterances in German and Polish.

1. Pragmalinguistischer Hintergrund

In *Principles of Pragmatics* bezeichnete Geoffrey Leech jene Interaktionen, in denen ein höherer inferentieller Aufwand (etwa die Rekonstruktion von Implikaturen, vgl. M. Haugh 2015) notwendig ist, um die Oberflächenstruktur (das Gesagte bzw. die Lokution) mit der Tiefenstruktur (das Gemeinte bzw. die Illokution) in Bezug zu setzen, als Fälle *obliquer* Kommunikation (G. Leech 1983: 97). Aus der Sicht der illokutionären Struktur haben wir es bei obliquer Kommunikation mit indirekten Sprechakten zu tun, in denen der Sprecher „zwar meint, was er sagt, aber darüber hinaus noch etwas mehr meint“ (J. R. Searle 1982: 51). Indirekte Sprechakte sind solche, „bei denen ein illokutionärer Akt indirekt, über den Vollzug eines andern, vollzogen wird“ (J. R. Searle 1982: 52). In indirekten Sprechakten haben wir zwei illokutionäre Rollen: eine primäre Illokution (was der kommunikativen Intention des Sprechers entspricht) und eine sekundäre Illokution (was sich aus der wörtlichen

Bedeutung des Satzes rekonstruieren lässt) (J. R. Searle 1982: 54). Schon Searle erkannte in diesen Sprechakten eine sehr breite linguistische Problemstellung: Wie versteht der Adressat den nicht-wörtlichen primären illokutionären Akt (d.h. was eigentlich gemeint ist) aufgrund dessen, dass er den wörtlichen sekundären illokutionären Akt (d.h. was gesagt wird) versteht?

Im folgenden Beitrag wird der Versuch unternommen, über die pragmalinguistischen Mechanismen hinaus auch performative Faktoren (phonetische Realisierung) zu untersuchen, die dazu führen, dass die primäre Illokution vom Sprecher und Hörer richtig ausgedrückt bzw. erkannt wird. Ironie und verbale Höflichkeit¹ sind „Paradebeispiele“ dafür, extreme Beispiele für oblique Kommunikation und indirekte Sprechakte sind Scheinbeleidigungen² (*mock impoliteness*, s. G. Leech 1983: 142ff., S. Bonacchi 2014) und Sarkasmus³ (*mock politeness*, s. J. Culpeper 1996: 356). Im Falle von Scheinbeleidigungen wird eine supportive Illokution durch eine derogative Äußerung (z.B. „du arsch!“ als Ausdruck der Bewunderung), im Falle von Sarkasmus eine feindliche bzw. aggressive Illokution durch eine supportive Äußerung (etwa ein Kompliment, z.B. „Tolle Arbeit!“ als ironisches Lob bzw. als Kritik für eine schwache Leistung) realisiert. Sowohl Sarkasmus als auch Scheinbeleidigungen sind Gegenstand zahlreicher linguistischer Untersuchungen gewesen⁴, bei denen vor allem auf die illokutionären Mechanismen und auf die besonderen Sprecher-Attitüden⁵ eingegangen wurde.⁶ Die Tatsache, dass diese dermaßen indirekten Äußerungen vom Rezipienten adäquat verstanden werden, wird auf unterschiedliche Art und Weise erklärt: durch einen erhöhten Kooperationsgrad zwischen Interaktanten (als Fälle der in-group-Kommunikation, vgl. V. Žegarac 1998 und J. Mateo/ F. Yus 2010), als Relevanzeffekt bzw. als „kontextueller Effekt“ (D. Sperber/ D. Wilson 2002: 113–115, D. Blakemore 2002, R. Carston 2006). Nach einer genaueren Analyse zeigt sich, dass bei korrekter Rekonstruktion der kommunikativen Intention des Sprechers nicht immer ein höherer inferentieller interpretatorischer Aufwand seitens des Interlokutors erforderlich ist, sondern dass diese Äußerungen oft ohne großen interpretatorischen Aufwand unmittelbar korrekt verstanden werden. Es wurde auch gezeigt, dass nicht nur der Kontext *sensu lato*, sondern auch die Modalitäten der Realisierung (vor allem Gestik, Mimik, Stimme) wesentlich dazu beitragen, dass die intendierte Bedeutung vom Interaktanten adäquat rekonstruiert wird (vgl. S. A. Stadler 2006, J. Culpeper 2011, S. McKinnon/ P. Prieto 2014). In diesen Studien über die multimodale Dimension dieser Sprechakte wurde gezeigt,

¹ Verbale Höflichkeit basiert prinzipiell auf obliquer bzw. indirekter Kommunikation, weil der Direktheit oft die Gefahr des Gesichtsverlustes (im Sinne E. Goffmans 1986) innewohnt.

² Die wörtliche Übersetzung von „*mock impoliteness*“ ist Scheinunhöflichkeit, allerdings wird hier der Ausdruck „Scheinbeleidigungen“ wegen des offensiven (und nicht so sehr unhöflichen) Charakters der pseudovokativen Formen (nach F. d’Avis/ J. Meibauer 2013) vorgezogen.

³ P. Brown/ S. Levinson (1987: 221) betrachten Sarkasmus als *off-record*-Strategie, die auf Ironie und Ambiguität basiert. Weitere Linguisten betrachten Sarkasmus als *mock politeness* (J. Culpeper 1996, J. Haiman 1998, D. S. Kaufer 1981, P. Rockwell 2000, 2006, J. Leggit/ R. Gibbs 2000, Wilson 2013) und weisen ihm eine aggressive Illokution zu. Zu den phonetischen Merkmalen von Sarkasmus vgl. H. S. Cheang/ M. D. Pell 2008 und C. Nakassis/ J. Snedeker 2002.

⁴ Für *mock impoliteness* vgl. G. Leech 1983: 142 ff., E. Nowik-Dziewicka 2012, J. Culpeper 2012, S. Bonacchi 2013 und 2014.

⁵ Für den Attitüde-Begriff vgl. H. Arndt Beckman/ R. W. Janney 1987: 139, 379f.

⁶ Für einen umfassenden Forschungsüberblick vgl. S. Bonacchi 2014

dass der jeweilige Adressat imstande ist, die vom Sprecher intendierte Bedeutung dank dessen „performative Leistung“ zu rekonstruieren bzw. zu verstehen. Es geht also nicht nur um die besondere Beziehung zwischen Sprecher und Adressat,⁷ sondern auch um ein von Sprecher und Adressat geteiltes Wissen darüber, welche Realisierungsbedingungen bzw. Erwartungen – eine entsprechende Gestik, Mimik, Kontextbedingungen, Prosodie – erfüllt werden müssen (Erwartungen), damit das Gemeinte trotz des Gesagten richtig rekonstruiert wird. Sowohl im Falle des Sarkasmus als auch im Falle der Scheinbeleidigungen kann es passieren, dass die Intention des Sprechers nicht erkannt wird und der indirekte Sprechakt misslingt, weil einige Merkmale im Merkmalsbündel fehlen – das ist zum Beispiel der Fall, wenn etwa ein Gruppenfremder versucht, Banter-Äußerungen zu realisieren, und dabei z.B. das Sprechtempo nicht richtig abstimmt und erwartete Dehnungen ausbleiben.

Im vorliegenden Beitrag werden die Ergebnisse einer Studie über die phonetischen Merkmale von Scheinbeleidigungen präsentiert, die im Rahmen des Forschungsprojektes MCCA (Multimodal Communication: Culturological Analysis) und in Zusammenarbeit des Instituts für Slavistik sowie des Instituts für Computerlinguistik und Phonetik der Universität des Saarlandes (Deutschland) und des Instituts für Fachkommunikation und interkulturelle Studien der Warschauer Universität durchgeführt wurde. Das Projekt MCCA verfolgte das Ziel, die kulturologischen Merkmale von (un)höflichen Äußerungen mit deren suprasegmentalen (phonetischen) Merkmalen in den untersuchten Sprachen Deutsch und Polnisch in Verbindung zu setzen. Höflichkeit ist zwar ein komplexes soziokulturelles Phänomen, aber aus phonetischer Sicht lässt sich der „höfliche Effekt“ meist mit Supportivität⁸ identifizieren. Während der Durchführung des Projektes wurde deutlich, dass sich angesichts der starken situationellen Kontextgebundenheit und der idiosynkratischen Variabilität der (un)höflichen Äußerungen keine einzelnen suprasegmentalen Merkmale (vor allem prosodische Merkmale) herausfinden lassen, die allein zum Zustandekommen des supportiven Effektes führen. Es sind eher Bündel von Merkmalen, die in enger Interdependenz auftreten und zum Zustandekommen des supportiven Effektes führen. Da die verbale Ebene eine sehr wichtige Rolle spielt und im Endeffekt zu einer extrem starken Variabilität führt, wurde ein Experimentdesign konzipiert, in dem bei konstanter lokutiver Ebene die phonetische Realisierung der Faktor war, der aus der Sicht der Perlokution zum Zustandekommen des supportiven bzw. derogativen Effektes führte. So wurden Äußerungen verglichen, die aus lokutiver Sicht identisch sind, aber durch bestimmte Realisierungsmerkmale entgegengesetzte Illokutionen – eine aus der Sicht der Höflichkeitstheorie gesichtsgefährdende bzw. derogative vs. eine gesichtsverstärkende bzw. supportive Illokution – ausdrücken (können). Ihr Effekt wurde mit muttersprachlichen Probanden untersucht.

⁷ Im Falle der Scheinbeleidigungen: „Ich darf so sprechen, weil wir Freunde sind“ (vgl. G. Leech 1983), im Falle des Sarkasmus als *off-record-Strategie* (P. Brown/ S. Levinson 1987: 221), die ermöglicht, direkte Sanktionen zu vermeiden.

⁸ Die vorliegende Studie wurde vom Polnischen Wissenschaftszentrum (Narodowe Centrum Nauki) im Rahmen des Projektes „Suprasegmentale und kulturologische Analyse von kommunikativen Interaktionen, die durch (Un)Höflichkeit markiert sind“ (kurz: MCCA, Multimodal Communication: Culturological Analysis, UMO-2012/04/M/HS2/00551) finanziell mitgetragen.

Es wurde dabei von der Annahme ausgegangen, dass (Un)Höflichkeitsakte als multimodale Akte aufzufassen sind, deren Kohärenz durch die Konvergenz der Informationen realisiert wird, die durch die verschiedenen Modalitäten vermittelt werden (R. Zwaan/ G. Radvansky 1998, D. Wilson/ T. Wharton 2006: 1559, S. McKinnon/ P. Prieto 2014).

2. Problemstellung

Die Theorie der universalen Aspekte der affektiven Interpretation der biologischen Codes nach C. Gussenhoven (2002) setzt voraus, dass Sprecher und Hörer in der Kommunikation von den universalen Aspekten der Prosodie Gebrauch machen. Demnach werde eine supportive (im Folgenden „positive“) bzw. eine derogative (im Folgenden „negative“) Sprecherattitüde grundsätzlich durch unterschiedliche prosodische Parameter realisiert bzw. signalisiert. C. Gussenhoven (2002: 47) nimmt drei biologische Codes an: den *frequency code*, den *effort code* und den *production code*. Der *production code* basiert auf der Tatsache, dass während der Ausatmungsphase ein gradueller Energieverlust stattfindet und das durchschnittliche F0-Niveau allmählich absinkt. Gussenhovens *effort code* besagt, dass die Sprecher ihren Produktionsaufwand je nach Wichtigkeit der Konstituenten permanent ändern. Wichtige Konstituenten werden mit größerem Aufwand (*effort*) und demnach mit einer präziseren Artikulation produziert als ihre angrenzenden weniger wichtigen Konstituenten, was sich im akustischen Signal in Exkursion der Tonhöhenbewegungen und einem global vergrößerten Tonhöhenumfang niederschlägt. Die affektive Interpretation des *effort code* könnte unter anderem den Eindruck der Zuvorkommenheit und der Hilfsbereitschaft wecken. Der *frequency code* (übernommen von J. J. Ohala 1983, 1984, 1994) bezieht sich auf die Korrelation der Größe des Kehlkopfs mit der Geschwindigkeit der Stimmlippenschwingungen. Der Zusammenhang zwischen der Grundfrequenz und der Körpergröße wurde in der nonverbalen Kommunikation vieler Spezies festgestellt (D. L. Bolinger 1964, E. S. Morton 1977). Größere Individuen haben größere Kehlköpfe. Die Stimmlippen in den größeren Kehlköpfen haben eine größere Masse und schwingen langsamer als die kürzeren und dünneren Stimmlippen in den kleineren Kehlköpfen. Deshalb produzieren größere Stimmlippen im Vergleich zu kleineren Stimmlippen eine tiefere Grundfrequenz. Dementsprechend wird eine tiefere Tonhöhe mit größeren Individuen assoziiert und eine höhere Tonhöhe mit kleineren Individuen. Die affektive Interpretation des *frequency code* ermöglicht es, Dominanzverhältnisse bzw. Machtverhältnisse anhand der prosodischen Realisierung zu deuten. Grundsätzlich gilt, dass je tiefer die Grundfrequenz eines Sprechers ist, desto dominanter und bedrohender er „klingen“ kann. Je höher die Grundfrequenz eines Sprechers ist, desto submissiver und unsicherer kann er „klingen“. Obwohl weitere Faktoren (wie idiosynkratische Variation und kulturelle Bedingtheit) zu Abweichungen von diesem Grundsatz führen können (vgl. dazu A. Chen et al. 2004, S. Grauwunder/ B. Winter 2010), liefert er die Basis für die Studien über die Prosodie der Höflichkeit und Unhöflichkeit (vgl. dazu u.a. S. McKinnon/ P. Prieto 2014).

Im Allgemeinen lässt sich feststellen, dass eine Interpretation des *frequency code* und des *effort code* in Bezug auf den *gesichtsverstärkenden* bzw. *gesichtsgefährdenden* Charakter von (Un)Höflichkeitsakten dazu führt, dass eine höhere Grundfrequenz und ein größerer Stimmtoumfang als Merkmale der supportiven bzw. höflichen, eine tiefere

Grundfrequenz und ein geringerer Stimmtonumfang als Merkmale der feindlichen/derogativen bzw. unhöflichen Sprechakte interpretiert werden können. Diese Annahme wird in der vorliegenden Studie anhand der Daten geprüft. Vor allem wird auf die Frage eingegangen, in welchem Grade die Grundfrequenz mit anderen prosodischen Parametern korreliert, um einen gegebenen Effekt beim Hörer auszulösen.

Die vorliegende Studie wird durch folgende Fragestellungen strukturiert:

- Was ist die prosodische Ausprägung der positiv/supportiv vs. negativ/derogativ intendierten Sprecherattitüde?
- Unterscheiden sich polnische und deutsche Sprecher/Hörer in der Ausnutzung der akustischen Parameter?
- Sind einzelne Sprecher innerhalb einer Sprache genauso unterschiedlich wie Sprecher verschiedener Sprachen?

3. Produktionsexperiment

3.1. Material, Versuchspersonen und Methode

Um zu untersuchen, ob die „supportive“ bzw. die „derogative“ Bedeutung einer Äußerung mit den verschiedenen akustischen Parametern des Sprachsignals im Polnischen und im Deutschen korreliert, wurden vier „offensive Äußerungen“ pro Sprache aufgenommen, die von den Sprechern als supportiv bzw. derogativ intendiert waren. Die vier Äußerungen waren mit fester verbaler Information (lexikalische Bedeutung) und offener vokaler Information (offene vokale Realisierung). Vier polnische und vier deutsche Muttersprachler (jeweils 2 Männer und 2 Frauen) wurden gebeten, die Äußerungen in der derogativen (negativen) und supportiven (positiven) Valenz zu realisieren. Das Durchschnittsalter der deutschen Sprecher war 29 Jahre und der Sprecherinnen 32 Jahre. Bei den polnischen Muttersprachlern war das Durchschnittsalter für die Sprecher 34,5 und für die Sprecherinnen 34.

Die vorgegebenen deutschen Testsätze waren die folgenden:

- (1) *Du bist ein Arsch!*
- (2) *Hey, Alter, was machst du denn hier!*
- (3) *Du hast es geschafft, du Sau!*
- (4) *Ach, du hast ja sowieso immer eine Eins, Du Penner!*

Die vorgegebenen polnischen Testsätze waren die folgenden:

- (1) *Ale z ciebie pies na baby!*
(dt. wörtliche Übersetzung: „Du bist aber ein Weiberhund!“)
- (2) *Ty draniu!*
(dt. wörtliche Übersetzung: „Du Schurke!“)
- (3) *Ty diable! Zniszczyłeś całą konkurencję!*
(dt. wörtliche Übersetzung: „Du Teufel! Du hast alle Konkurrenten vernichtet!“)
- (4) *Ale ty zaliczasz te laseczki, ty złamasie!*
(dt. wörtliche Übersetzung: „Du kriegst alle Mädels rum, du Invalidenschwanz!“)⁹

⁹ In den polnischen Testsätzen werden beleidigende Pseudovokative (vgl. D’Avis/ Meibauer 2014) benutzt: „pies na baby“ (wörtlich: „Weiberhund“, Beleidigung durch Animalisierung), „ty draniu“

Die Äußerungen wurden mehrmals als Reaktion auf einen von dem Versuchsleiter situativ beschriebenen Kontext produziert. Der Kontext sollte den SprecherInnen die Aufgabe erleichtern, sich je eine Situation vorzustellen, in der sie die jeweilige Äußerung als freundlich (kumpelhaft supportiv) bzw. unfreundlich (feindlich/aggressiv) benutzen würden. Danach hörte sich jede(r) SprecherIn alle Produktionen an und entschied, welche die beste (die „am meisten gelungene“) Realisierung war. Diese wurde akustisch analysiert und im Perceptionsexperiment benutzt. Insgesamt wurden 32 Äußerungen pro Sprache (4 SprecherInnen x 4 Äußerungen x 2 Valenzen) ausgewertet.

3.2. Akustische Analyse

Eine Vielzahl an Untersuchungen hat gezeigt, dass Messungen, bezogen auf Grundfrequenz, Intensität und Sprechtempo in Bezug auf die supportive bzw. derogative Sprecherattitüde unterschiedlich sein können (I. R. Murray/ J. L. Arnott 1993, S, McKinnon/ P. Prieto 2014). In unserer Studie wurden daher folgende akustische Parameter mithilfe des Programms Praat abgeleitet:

- a) durchschnittliche Grundfrequenz für die Gesamtäußerung – gemessen nach Ausschluss der oberen und unteren 5,5%, um Ausreißer wegen Oktavsprüngen zu vermeiden;
- a) Tonhöhenumfang (gemessen in Halbtönen) – nach Ausschluss der oberen und unteren 5,5% (um Ausreißer wegen Oktavsprüngen zu vermeiden) wird der tiefste und der höchste F0-Wert pro Äußerung ermittelt. Die Differenz zwischen diesen beiden Messpunkten ergibt den Tonhöhenumfang für die jeweilige Äußerung;
- b) Durchschnittliche Intensität für die Gesamtäußerung;
- c) Intensitätsumfang (gemessen in Dezibel) – pro Äußerung wird der tiefste und der höchste Intensitätswert ermittelt. Die Differenz zwischen diesen beiden Messpunkten ergibt den Intensitätsumfang für die jeweilige Äußerung;
- d) durchschnittliche Silbendauer (gemessen in Millisekunden) – wird durch Division der Dauer der gesamten Äußerung durch die Anzahl der Silben pro Äußerung berechnet.

Um die Effekte der jeweiligen Sprecherattitüde auf die einzelnen akustischen Parameter zu untersuchen und eventuelle Unterschiede zwischen den Sprachen festzustellen, wurden Lineare Gemischte Modelle (LMM) mit dem jeweiligen Wert des Parameters als abhängige Variable, Satz als Zufallsfaktor, Attitüde (freundlich bzw. positiv / unfreundlich bzw. negativ) und Sprache (Deutsch/Polnisch) als feste Faktoren sowie ihre möglichen Wechselwirkungen gerechnet. Tukey Post-hoc Tests, soweit geeignet, wurden ebenfalls durchgeführt.

Die Werte der F0-bezogenen Parameter (durchschnittliche Grundfrequenz und Stimmumfang) unterscheiden sich in den beiden Sprachen und in den beiden Attitüden nicht signifikant voneinander (s. Abb. 1 und Abb. 2). Die Daten zeigen allerdings eine Tendenz zur Erhöhung des Stimmumfangs in der positiven Attitüde in beiden Sprachen.

(wörtlich: „du Schurke“), „ty diable“ (wörtlich: „du Teufel“, Beleidung durch Verletzung des religiösen Tabus), „ty zlamasie“ (wörtlich: „Du gebrochener (Schwanz)“ (Beleidung durch Verletzung des sexuellen Tabus). Für eine breitere Analyse der Offensivität vgl. S. Bonacchi 2012.

Sprache und Attitüde haben einen Haupteffekt auf die durchschnittliche Intensität. Die Gruppe der polnischen SprecherInnen spricht signifikant lauter als die Gruppe der deutschen SprecherInnen ($F [1, 9,222] = 108,8648, p < 0,001$), und die Äußerungen in der derogativen Intention (negative Sprecherattitüde) über die Sprachen hinweg sind signifikant lauter als die in der supportiven Intention (positive Sprecherattitüde) ($F [1, 48] = 8,6758, p < 0,01$). Die Analyse zeigt auch eine signifikante Interaktion zwischen Sprache und Attitüde ($F [1, 48] = 28,4029, p < 0,001$). Die polnischen SprecherInnen sind um 4,27 dB lauter in der negativen Attitüde. Die deutschen SprecherInnen sprechen hingegen gleich laut in beiden Bedingungen (s. Abb. 3). Die SprecherInnen beider Gruppen benutzen einen größeren Intensitätsumfang in der negativen Attitüde ($F [1, 48] = 5,9529, p < 0,05$) – s. Abb. 4).

Die Bar Plots in Abb. 5 zeigen die durchschnittliche Silbendauer im Deutschen und im Polnischen in der negativen und in der positiven Attitüde. Die statistische Analyse zeigt einen signifikanten Effekt der Attitüde ($F [1, 48] = 7,6198, p < 0,01$) sowie eine Interaktion zwischen Attitüde und Sprache ($F [1, 48] = 8,9792, p < 0,01$). Die deutschen SprecherInnen sprechen schneller in der positiven Attitüde. Die durchschnittliche Silbendauer in der negativen Attitüde ist um 33,06 Millisekunden länger im Vergleich zur positiven Attitüde. Polnische SprecherInnen zeigen diesen Unterschied nicht.

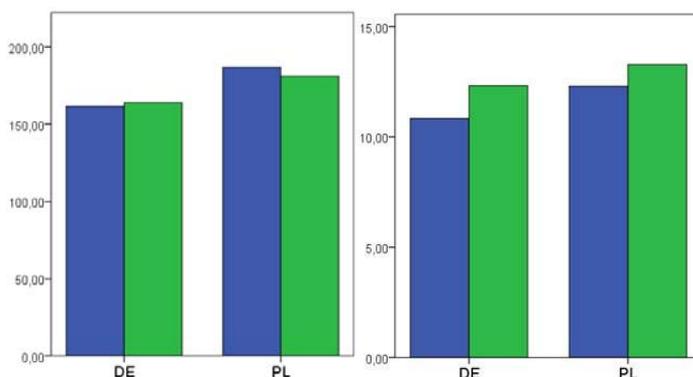


Abbildung 1. Grundfrequenz (F0)

Abbildung 2. F0-Umfang

(blau: positiv, grün – negativ)

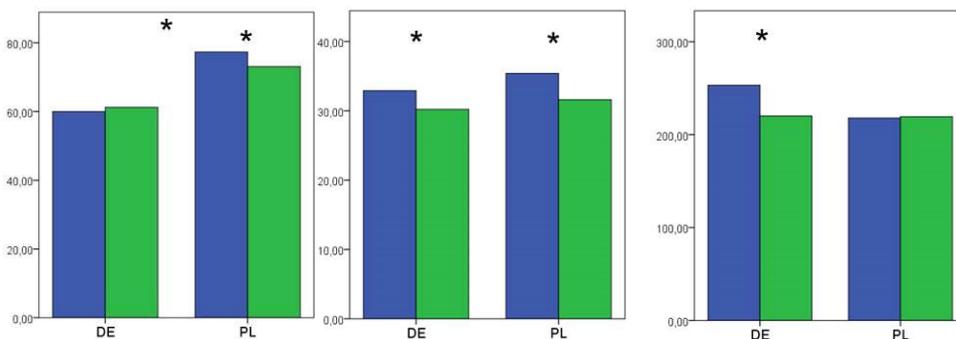


Abbildung 3. Intensität

Abbildung 4. Intensitätsumfang

Abbildung 5. Dauer

3.3. Nukleare Tonakzente

Zusätzlich wurden die nuklearen Tonakzente in jeder Äußerung nach der ToBI-Konvention etikettiert (G-ToBI für die deutschen (vgl. M. Grice/ S. Baumann 2002 und M. Grice, S. Baumann/ R. Benz Müller 2005) und eine angepasste ToBI-Version für die polnischen Daten). Die verwendete Notationskonvention stammt aus dem autosegmental-metrischen Ansatz der Intonationsbeschreibung, die sich in Anlehnung an die Arbeiten von J. Pierrehumbert (1980), M. E. Beckman/ J. Pierrehumbert (1986) und J. Pierrehumbert/ M. E. Beckman (1988) entwickelt hat. Dieser Ansatz verwendet zwei Töne: H(igh)-Töne, die als Gipfel, und L(ow)-Töne, die als Täler im Grundfrequenzverlauf realisiert werden. Diese Töne sind an bestimmte Punkte in einer Äußerung geknüpft und haben keine absoluten Werte, sondern sind auf der phonologischen Beschreibungsebene relativ. Es gibt monotonale Tonakzente (z.B. H* oder L*) und bitonale Tonakzente (z.B. L*+H, H+L*). Der Ton mit dem Stern wird mit der akzentuierten Silbe assoziiert, d.h. gesternte Töne werden als Gipfel oder Täler auf der akzentuierten Silbe realisiert. Ungesternte Töne in einem Tonakzent werden vom gesternten Ton durch einen gewissen Zeitabstand getrennt. ‘!’ kennzeichnet die Herabstufung eines H-Tons (*downstep*) in Relation zu einem vorangehenden H-Ton. ‘^’ kennzeichnet die Erweiterung des Stimmumfangs bei einem H-Ton (*upstep*) in Relation zu einem vorangehenden H-Ton. ‘<’ und ‘>’ kennzeichnen jeweils einen späten und einen frühen Gipfel in Relation zu den Grenzen der Akzentsilbe.

3.3.1. Nukleare Tonakzente im Deutschen

Die deutschen SprecherInnen benutzen folgende Tonakzente H*, L*, H+L*, L+H*, H+!H* (s. Tabelle 1). H* ist der unmarkierte Akzent des deutschen Toninventars (s. Abb. 8). Er zeichnet sich durch einen Tonhöhengipfel aus, der meistens gegen das Ende der Akzentsilbe erreicht wird.

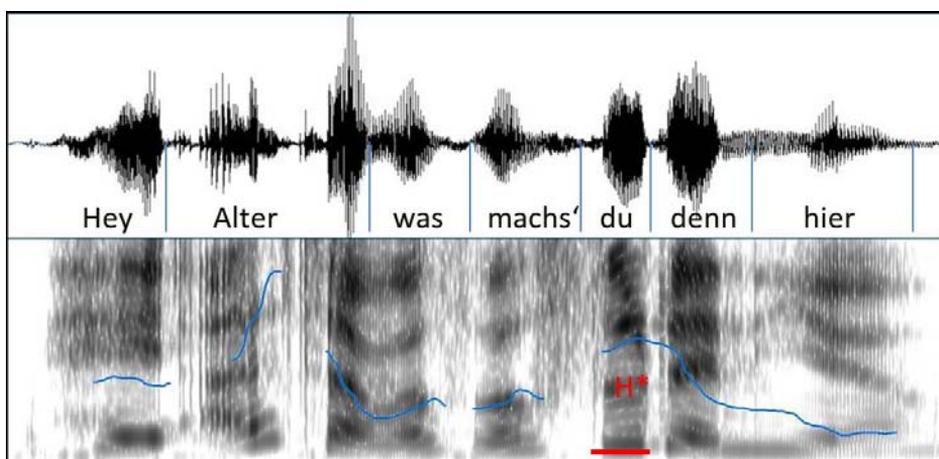


Abbildung. 8. „Hey, Alter, was machst du denn hier!“, realisiert von der Sprecherin JJ in der negativen Sprecherattitüde (Tonakzent H*)

Bei dem L*-Tonakzent wird ein Tal im F0-Verlauf erreicht. Er zeichnet sich durch eine tiefe Stimmlage im unteren Drittel des Sprechstimmumfangs aus. Bei L+H* findet ein steiler Anstieg auf der akzentuierten Silbe statt, und der Tonhöhengipfel wird erst spät in der akzentuierten Silbe (manchmal auch danach) erreicht (s. Abb. 9). Auf der Silbe vor dem Akzent (oder zu Beginn der Akzentsilbe) wird ein Tiefpunkt erreicht - die Kontur fällt bis dahin also ab, falls genügend Silben vor dem Akzent zur Verfügung stehen. Oft ist der Stimmumfang erweitert. Der bitonale Akzenttyp H+!H* zeichnet sich durch einen Fall auf der akzentuierten Silbe von einer hohen in eine mittlere Stimmlage aus. Die vorangehende Silbe ist höher. Bei H+L* ist die akzentuierte Silbe tief. Der Tonhöhen-Gipfel befindet sich unmittelbar vor der Akzentsilbe. Bei den Akzenttypen H* und L+H* wird die akzentuierte Silbe als hoch oder steigend wahrgenommen, bei H+!H* als hoch-fallend, bei L* und H+L* als tief. Zusätzlich entsteht bei H+L* der Eindruck eines großen Tonhöhen sprungs in eine tiefe Stimmlage.

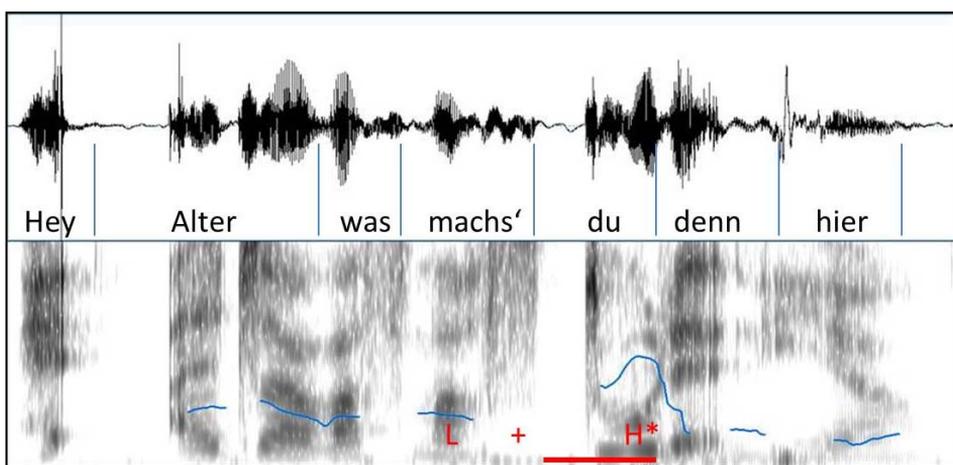


Abbildung 9. „Hey, Alter, was machst du denn hier!“, realisiert von der Sprecherin JJ in der positiven Sprecherattitüde (Tonakzent L+H*)

In Tabelle 1 sind die realisierten nuklearen Tonakzente pro SprecherIn, Testsatz und Bedingung für die deutschen Daten angegeben. Es fällt auf, dass sowohl die hohen/steigenden Tonakzente als auch die tiefen/fallenden Tonakzente gleichmäßig in den positiven und negativen Bedingungen verteilt sind. Die SprecherInnen haben in der positiven/freundlichen Attitüde 11 Mal (10 H* und 1 L+H*) und in der negativen/unfreundlichen Attitüde ebenfalls 11 Mal (7 H* und 4 L+H*) hohe/steigende nukleare Akzente realisiert. Tiefe/fallende nukleare Tonakzente wurden 5 Mal in den freundlichen Realisierungen (1 L*, 2 H+!H* und 2 H+L*) und 5 Mal in den unfreundlichen Realisierungen festgestellt (1 L*, 3 H+!H* und 1 H+L*).

SprecherInnen	Satz 1		Satz 2		Satz 3		Satz 4	
	positiv	negativ	positiv	Negativ	positiv	Negativ	positiv	negativ
FZ (m)	H* (du)	H* (du)	H* (du)	H* (du)	H+!H* (Penner)	H+!H* (Penner)	H+L* (Sau)	H+!H* (Sau)
RM (m)	L+H* (du)	L+H* (machst)	H* (A:rsch)	L+H* (Arsch)	L* (Penner)	H* (Penner)	H+L* (Sau)	H* (Sau)
SK (m)	H* (du)	L+H* (du)	H* (du)	H* (Arsch)	H* (Penner)	H* (Penner)	H* (Sau)	H+!H* (Sau)
JJ (m)	H* (du)	L+H* (du:)	H* (du)	H* (A:rsch)	H+!H* (Penner)	H+L* (Penner)	H* (Sau)	L* (Sau)

Tabelle 1. Nukleare Tonakzente, realisiert von den deutschen SprecherInnen pro Bedingung und Äußerung

Aus der Tabelle wird ersichtlich, dass im Falle des Testsatzes 2 die beiden SprecherInnen genauso wie die Sprecher hohe/steigende nukleare Tonakzente realisieren, aber im Unterschied zu den Sprechern disambiguieren sie zwischen der freundlichen und unfreundlichen Attitüde, indem sie den Nuklearakzent in der freundlichen Bedingung auf dem Wort ‚du‘ und in der unfreundlichen Bedingung auf dem Wort ‚Arsch‘ realisieren. Auffällig ist der männliche Sprecher FZ, der bei Testsätzen 1, 2 und 3 in den beiden Bedingungen den gleichen nuklearen Akzent benutzt.

3.3.2. Nukleare Tonakzente im Polnischen

Die von den polnischen SprecherInnen benutzten nuklearen Tonakzente sind H*, \wedge H*, !H*, L+H*, L+!H*, H+!H* und L* (s. Tabelle 2). Der H*-Akzentton zeichnet sich durch einen Tonhöhengipfel aus, der meistens am Anfang der Akzentsilbe erreicht wird. Die akzentuierte Silbe hört sich hoch an. Eine Tonhöhenmodifikation von H* ist durch Downstep oder Upstep möglich (!H*, \wedge H*). Der L*-Tonakzent zeichnet sich durch eine tiefe Stimmlage im unteren Drittel des Sprechstimmumfangs aus. Die akzentuierte Silbe hört sich tief an. Bei einem L+H* wird unmittelbar vor oder in der akzentuierten Silbe ein tiefer Zielpunkt erreicht, danach folgt ein Anstieg. Der Tonhöhengipfel wird erst spät in der akzentuierten Silbe erreicht. Oft ist die Silbe gedehnt (s. Abb. 10). Bei diesem Tonakzent ist ebenfalls eine Tonhöhenmodifikation durch den Downstep möglich (L+!H*).

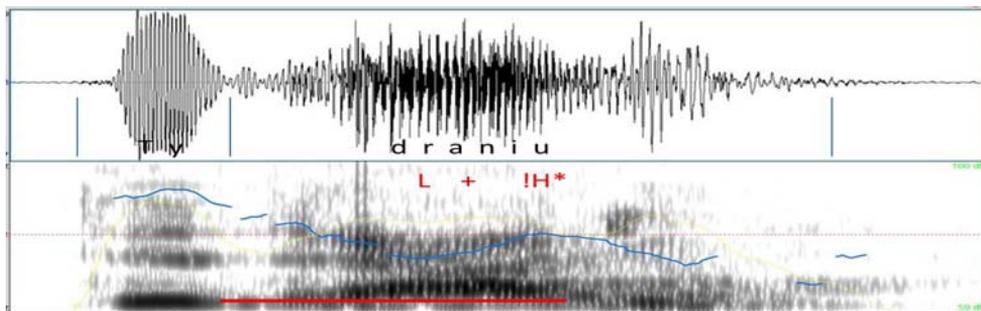


Abbildung 10. „Ty draniu“ (Du Schurke), realisiert vom männlichen Sprecher RG in der positiven Sprecherattitüde (Tonakzent L+!H*)

Der bitonale Akzenttyp H+!H* zeichnet sich durch eine mittlere Stimmlage der akzentuierten Silbe aus. Der Tonhöhen-Gipfel befindet sich unmittelbar vor der Akzent-silbe (s. Abb. 11). Die Gipfelakzente kann man generell in 2 Gruppen aufteilen: Tonakzente mit frühem Gipfel (H*, !H*, ^H*, H+!H*) und Tonakzente mit spätem Gipfel (L+H*, L+!H*).

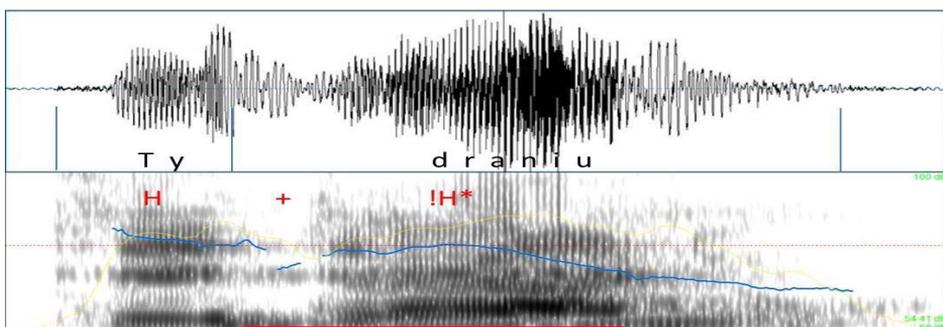


Abbildung. 11. „Ty draniu“ (Du Schurke), realisiert vom Sprecher RG in der negativen Sprecherattitüde (Tonakzent H+!H*)

In Tabelle 2 sind die realisierten nuklearen Tonakzente pro SprecherInnen, Testsatz und Bedingung für die polnischen Daten angegeben.

SprecherInnen	Satz 1		Satz 2		Satz 3		Satz 4	
	positiv	negativ	positiv	Negativ	positiv	negativ	positiv	negativ
RG (m)	L* H- H% (baby)	H* (pies)	L+!H* (dranju)	H+!H* (dranju)	L+!H* (di- able)	^H* (di- able)	L* H- H% (złama- sie)	H+!H* (złama- sie)
AD (m)	L+!H* (baby)	^H* (baby)	L+H* (dranju)	H* (dranju)	L+H* (di- able)	^H* (di- able)	!H* (złama- sie)	H+!H* (złama- sie)
PM (m)	!H* (baby)	H* (pies)	L+!H* (dranju)	L+H* (dranju)	L+H* (di- able)	^H* (di- able)	L+H* (ty)	H+!H* (złama- sie)

LG (m)	!H* (pies)	H* (pies)	L+!H* (dranju)	H* (dranju)	L+!H* (di- able)	^H* (di- able)	H+!H* (złama- sie)	H+!H* (złama- sie)
--------	---------------	--------------	-------------------	----------------	------------------------	----------------------	--------------------------	--------------------------

Tabelle 2. Nukleare Tonakzente, realisiert von den polnischen SprecherInnen pro Bedingung und Äußerung

Bei genauerer Betrachtung der Tabelle wird ersichtlich, dass im Unterschied zu den deutschen SprecherInnen die polnischen SprecherInnen eine deutliche Präferenz für späte Gipfel in der freundlichen Attitüde und für frühe Gipfel in der unfreundlichen Attitüde haben. In der freundlichen Attitüde realisieren die SprecherInnen 10 Mal Akzente mit spätem Gipfel (4 L+H* und 6 L+!H*) und 4 Mal mit frühem Gipfel (3 !H* und 1 H+!H*). In der unfreundlichen Attitüde wurden 15 Mal nukleare Tonakzente mit frühem Gipfel produziert (5 H*, 5 ^H* und 5 H+!H*). Nur Sprecherin PM realisiert in dieser Bedingung einmal einen L+H* nuklearen Akzentton. Sprecher RG benutzt in der freundlichen Bedingung zweimal den tiefen L*, aber seine nukleare Intonationskontur ist steigend. Nach der tiefen akzentuierten Silbe steigt die Tonhöhe bis zum Ende der Äußerung hin an.

4. Perzeptionsexperiment

4.1. Material, Versuchspersonen und Methode

Die im Produktionsexperiment realisierten 32 Äußerungen pro Sprache (4 SprecherInnen x 4 Äußerungen x 2 Attitüden) dienten als Stimuli für das Perzeptionsexperiment. 49 polnische Versuchspersonen (Durchschnittsalter: 20,22; Standardabweichung: 4,83) haben die polnischen Stimuli und 29 deutsche Versuchspersonen (Durchschnittsalter: 28,97; Standardabweichung: 7,37) haben die deutschen Stimuli beurteilt. Alle Versuchspersonen waren Muttersprachler ohne Beeinträchtigung der Hörfunktion. Sie hatten die Möglichkeit, die Äußerungen so oft zu hören, wie sie wollten. Die Reihenfolge der Äußerungen wurde randomisiert, um eine Verzerrung der Ergebnisse zu vermeiden. Das Ausfüllen jedes online-Fragebogens nahm ca. 10-15 Minuten in Anspruch.

Abbildung 14. Online-Fragebogen

Das Perzeptionsexperiment wurde mittels der Internet-Plattform „Sosci-Survey“ (<https://www.soscisurvey.de/>) durchgeführt. Die polnischen und deutschen Teilnehmer an der online-Befragung wurden gebeten, sich die 32 Stimuli anzuhören und auf einer kontinuierlichen Skala anzugeben, wie „freundlich“ oder „unfreundlich“¹⁰ die Äußerung gemeint war (s. Abb. 14).

Pro Sprache wurden Lineare Gemischte Modelle mit Urteil (z-transformiert) als abhängige Variable, Stimulus und Versuchsperson als Zufallsfaktor, Attitüde (freundlich/unfreundlich) und SprecherInnen (4 pro Sprache) als feste Faktoren sowie alle ihre mögliche Wechselwirkungen gerechnet. Tukey Post-hoc Tests, soweit geeignet, wurden ebenfalls durchgeführt.

4.2. Ergebnisse: Deutsch

Die statistische Analyse der Daten zeigt einen Haupteffekt bezüglich der Attitüde ($F [1, 24] = 24,9708, p < 0,001$). Das bedeutet, dass die deutschen Versuchspersonen die Attitüden in den Stimuli signifikant unterschiedlich beurteilt haben. Der Mittelwert der Urteile für die positive Attitüde ist 0,45 und für die negative -0,45. Jedoch wurde eine Interaktion zwischen Attitüde und SprecherInnen ($F [1, 24] = 4,9621, p < 0,01$) gefunden. Diese Interaktion hat eine Auswirkung auf die Interpretation des Haupteffekts und bedeutet, dass die intendierte Attitüde nicht bei allen SprecherInnen erkannt wurde. Tabelle 4 enthält die Mittelwerte der Urteile für die Stimuli, aufgeteilt nach SprecherInnen und Attitüde, sowie die Ergebnisse der t-tests. Daraus wird ersichtlich, dass die Versuchspersonen nicht imstande waren, zwischen den intendierten Attitüden der männlichen Sprecher FZ und RM zu differenzieren.

FZ (m)		RM (m)		SK (f)		JJ (f)	
positiv	negativ	positiv	Negativ	positiv	negativ	positiv	negativ
-0,05	0,03	0,39	-0,37	0,47	-0,57	0,98	-0,87
nicht signifikant		nicht signifikant		p<0,001		p<0,001	

Tabelle 4. Mittelwerte der Urteile pro SprecherInnen und Attitüde

Die Balkendiagramme in Abb. 15 veranschaulichen die Ausnutzung des F0-Umfangs, des Intensitätsumfangs und der Dauer der vier deutschen SprecherInnen. Die SprecherInnen benutzen systematisch das Sprechtempo und mit Ausnahme von Sprecherin JJ den Intensitätsumfang, um die (Schein-)Beleidigung zu signalisieren – sie benutzen ein langsames Sprechtempo und einen größeren Stimmumfang, wenn sie eine unfreundliche Attitüde intendieren. In Bezug auf den Stimmumfang wenden die männlichen und die weiblichen SprecherInnen unterschiedliche Strategien an. Wie im Pkt. 2

¹⁰ Bei „freundlichen“ im Sinne von „supportiv“ und „unfreundlich“ im Sinne von „derogativ“ wurde auf lebensweltliche Bezeichnungen zurückgegriffen, damit die Probanden, die kein linguistisches Fachwissen hatten, die Anweisungen korrekt verstehen konnten. Die Teilnehmer basierten ihre Urteile auf ihrem intuitiven Wissen, ihnen wurden keine Definitionen von „freundlich“ und „unfreundlich“ gegeben.

beschrieben, liegt die symbolische Natur der Tonhöhe im Sinne des *frequency code* (J. J. Ohala 1983, 1984, 1994, C. Gussenhoven 2002) primär auf der Andeutung von Stärkerelationen. Im Sinne des *effort code* (C. Gussenhoven 2002) ist der Stimmumfang auf Entgegenkommen zurückzuführen. Während Sprecherinnen SK und JJ einen größeren Stimmumfang in der freundlichen Attitüde benutzen, reduziert Sprecher RM den F0-Umfang in dieser Bedingung, und Sprecher FZ benutzt den gleichen Stimmumfang in beiden Attitüden. Diese widersprüchlichen akustischen *Cues* behindern höchstwahrscheinlich die Versuchspersonen bei der Erkennung der intendierten Sprechereinstellung.

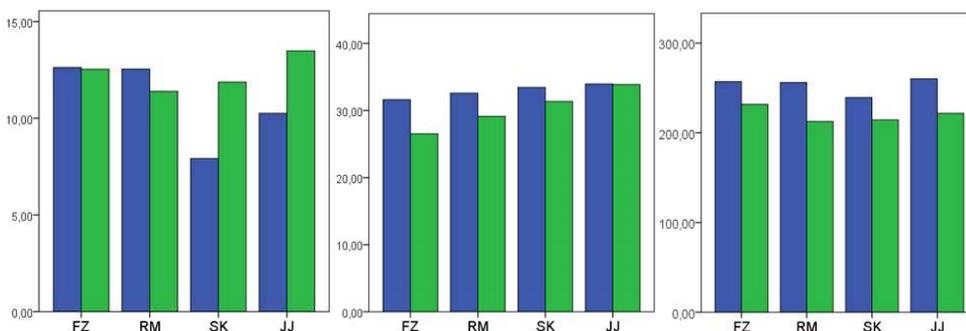


Abbildung 16. F0-Umfang (links), Intensitätsumfang (in der Mitte) und mittlere Silbendauer pro SprecherInnen in der positiven (grün) und negativen (blau) Bedingung (Deutsch).

4.3. Ergebnisse: Polnisch

Ein Haupteffekt bezüglich der Attitüde ($F [1, 24] = 69,9345, p < 0,001$) und der Sprecher ($F [3, 24] = 7,8547, p < 0,001$) tritt auf. Die Versuchspersonen haben die Attitüden der Stimuli signifikant unterschiedlich beurteilt. Die Stimuli mit der positiven Attitüde wurden im Schnitt mit dem Wert 0,50 beurteilt, die Stimuli mit der negativen Attitüde mit -0,50. Tabelle 3 enthält die Mittelwerte der Urteile für die Stimuli, aufgeteilt nach Sprecher und Attitüde. Daraus wird ersichtlich, dass die negative Attitüde von Sprecher AD und die positive Attitüde von Sprecherin PM nicht überzeugend erkannt werden. Beide Werte sind in der Nähe des Mittelpunktes der Skala 0. Es verhält sich genau umgekehrt mit der Erkennung der positiven Attitüde bei Sprecher AD 0,90 und der negativen Attitüde bei Sprecherin PM -0,83. Bei der auditiven Beurteilung der realisierten Äußerungen wurde festgestellt, dass bei der Realisierung aller supportiven Äußerungen Sprecher AD gelächelt hat. Sprecherin PM hat eine auffällig veränderliche Stimmqualität (behaucht, hart, rau).

RG (m)		AD (m)		PM (f)		LG (f)	
positiv	negativ	positiv	Negativ	positiv	negativ	positiv	negativ
0,55	-0,41	0,90	-0,06	0,11	-0,83	0,43	-0,71

Tabelle 3. Mittelwerte der Urteile pro SprecherIn und Attitüde

Die Balkendiagramme in Abb. 15 veranschaulichen die Ausnutzung des F0-Umfangs, des Intensitätsumfangs und der Dauer der vier polnischen SprecherInnen. Es fällt auf, dass die Sprecher die akustischen Parameter zum Signalisieren der jeweiligen Sprechereinstellung sehr unterschiedlich einsetzen. Sprecher RG und Sprecherin LG haben einen höheren Stimmumfang in der freundlichen Attitüde, Sprecher AD und Sprecherin PM in der unfreundlichen Attitüde. Was das Sprechtempo anbelangt, unterscheiden sich die SprecherInnen in der Wahl ihrer Strategie extrem voneinander. Im Vergleich zur positiven Attitüde ist Sprecher RM langsamer und Sprecher AD und Sprecherin LG sind schneller in der negativen Attitüde. Sprecherin PM macht keinen Gebrauch vom Sprechtempo zur Kennzeichnung der intendierten Sprechereinstellung. Darüber hinaus ist sie in den beiden Bedingungen langsamer als der Rest der Sprecher. Nur in Bezug auf den Intensitätsumfang zeigen die Sprecher eine gewisse Systematik: abgesehen von Sprecherin AD haben sie einen größeren Intensitätsumfang in der unfreundlichen Attitüde.

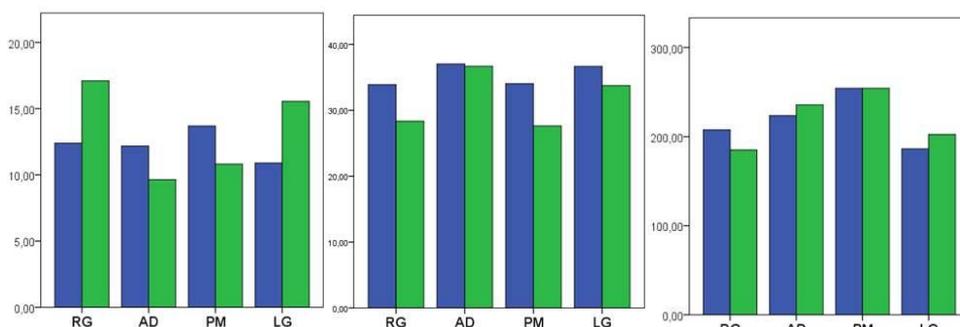


Abbildung 15. F0-Umfang (links), Intensitätsumfang (in der Mitte) und mittlere Silbendauer pro Sprecher in der positiven (grün) und negativen (blau) Bedingung (Polnisch)

5. Schlussfolgerungen

Unsere Studie hat unterschiedliche phonetische Profile bei der Produktion von supportiven und derogativen Äußerungen im Deutschen und im Polnischen gezeigt. Das Gelingen des indirekten Sprechaktes der Scheinbeleidigung (*mock impoliteness*) in den untersuchten Sprachen (verstanden als Anerkennung durch den Hörer der supportiven Intention des Sprechers trotz des Gebrauchs einer derogativen Lokution) hängt nicht nur von kontextuellen Faktoren und von der adäquaten Rekonstruktion von Implikaturen ab, sondern ist auch an „performative“ Faktoren („wie“ etwas gesagt wird) gebunden. Unter den phonetischen Merkmalen, die für das Gelingen dieses Sprechaktes relevant sind, wurden in der vorliegenden Studie die durchschnittliche Grundfrequenz, der Tonhöhenumfang, die durchschnittliche Intensität, der Intensitätsumfang, die durchschnittliche Silbendauer und die nuklearen Tonakzente untersucht.

Die Ergebnisse des Produktionsexperimentes haben gezeigt, dass bei der Realisierung solcher Sprechakte SprecherInnen auf ein Instrumentarium zurückgreifen, bestehend einerseits aus kontinuierlichen prosodisch-phonetischen Parametern und andererseits aus kategorialen intonatorischen Merkmalen, die in unterschiedlichen Graden der Kombination eingesetzt werden können. Das Perzeptionsexperiment hat gezeigt, dass

aus der Sicht des perlokutionären Effektes die Wahl der eingesetzten Mittel eine wichtige Rolle bei der Erkennung der intendierten Attitüde spielt. In beiden Sprachen hat sich folgendes herausgestellt: Wenn die SprecherInnen bei der Produktion der jeweiligen Attitüden die entsprechenden relevanten phonetischen und intonatorischen Merkmale nicht benutzt (vor allem Grundfrequenz) und dieses auch nicht durch eine ausbalancierende Verwendung von anderen Merkmalen (z.B. Intensität oder Sprechtempo) kompensiert haben, wurden diese Attitüden von den HörerInnen nicht (eindeutig) erkannt.

In Bezug auf die anfangs definierte Problemstellung lässt sich das Ergebnis wie folgt zusammenfassen:

- Die negative oder positive Valenz von Äußerungen kann auf der Grundlage von graduellen phonetischen und kategorialen intonatorischen Merkmalen auch ohne kontextuelle Informationen korrekt rekonstruiert werden (signifikanter perlokutiver Effekt, vgl. auch J. House 2008)¹¹.
- Bei der Produktion von *mock impoliteness* geht es um ein Merkmalsbündel, bei dem Tonakzent, Grundfrequenz, Intensität, Dauer, Stimmqualität (aber auch nichtverbale Vokalisierungen: Lächeln/Lacher) die wichtigsten Komponenten sind. Die vorliegende Studie zeigt, dass zwar die Grundfrequenz ein wichtiger Faktor ist, der zum Gelingen des Sprechaktes der *mock impoliteness* (supportive Tiefenstruktur mit derogativer Oberflächenstruktur) führt, allerdings ist dieser Faktor immer mit anderen phonetischen Faktoren verbunden, und nur das daraus resultierende Bündel ist entscheidend dafür, ob der Sprechakt gelingt oder nicht. Dieses Ergebnis relativiert die Aussagekraft des *frequency code* und des *effort code*.

Bei den untersuchten Produktionen der polnischen und deutschen SprecherInnen lässt sich folgendes feststellen:

- Bei den polnischen und den deutschen SprecherInnen ist ein größerer Intensitätsumfang in der negativen Attitüde festzustellen.
- Bei den deutschen SprecherInnen sind größere Dauerwerte in der negativen Attitüde festzustellen.
- Bei den polnischen SprecherInnen ist in beiden Attitüden eine größere Durchschnittsintensität als bei den Deutschen festzustellen.
- Die polnischen und die deutschen SprecherInnen verwenden unterschiedliche Akzenttypen. Die polnischen SprecherInnen zeigen eine deutliche Präferenz für späte Gipfel in der freundlichen Attitüde und für frühe Gipfel in der unfreundlichen Attitüde.

Die Ergebnisse des Produktions- und Perceptionsexperiments haben gezeigt, dass in beiden Sprachen eine relativ hohe Intersprecher-Variabilität festzustellen ist. Es soll schließlich betont werden, dass die Ergebnisse dieser Studie als qualitativ zu betrachten sind und in weiteren quantitativen Studien geprüft und durch weitere Variablen erweitert werden sollen.

¹¹ Dies relativiert die Kontext-Hypothese, vgl. D. Wilson/ T. Wharton 2006: 1559: „First, prosodic inputs to the comprehension process range from the ‘natural’ (e.g. an angry, friendly or agitated tone of voice) to the properly linguistic (e.g. lexical stress or lexical tone). Second, the effects of prosody are highly context-dependent: prosodic information interacts with information from many other sources during the comprehension process, and the same prosodic input may have different effects on different occasions.”

Bibliographie

- Arndt, H./ R. W. Janney (1985), *Politeness revisited: cross modal supportive strategies*. In: International Review of Applied Linguistics in Language Teaching 23 (4), 281–300.
- Arndt, H./ Janney, R. W. (1987), *Intergrammar: Toward an Integrative Model of Verbal, Prosodic and Kinesic Choices in Speech*. Berlin.
- Beckman, M.E./ J.B. Pierrehumbert (1986), *Intonational structure in English and Japanese*. In: Phonology Yearbook 3, 255–310.
- Blakemore, D. (2002), *Relevance and Linguistic Meaning*. Cambridge.
- Bolinger, D. L. (1964), *Intonation across languages*. In: J.H. Greenberg/ C.A. Ferguson/ E.A. Moravcsik (Hrsg.), *Universals of Human Language*, Vol. 2: Phonology. Stanford, CA., 471–524.
- Bonacchi, S. (2012), *Zu den idio-kulturellen und polykulturellen Bedingungen von aggressiven Äußerungen im Vergleich Polnisch-Deutsch-Italienisch*. In: M. Olpińska-Szkielko/ S. Grucza/ Z. Berdychowska/ J. Żmudzki (Hrsg.), *Der Mensch und seine Sprachen*. Frankfurt a.M. etc., 130–148
- Bonacchi, S. (2013), *(Un)Höflichkeit. Eine kulturologische Analyse Deutsch – Italienisch – Polnisch*. Frankfurt a.M.
- Bonacchi, S. (2014), *Scheinbeleidigungen und perfide Komplimente: kulturologische Bemerkungen zur obliquen Kommunikation in interkultureller Perspektive*. In: K. Lukas/ I. Olszewska (Hrsg.), *Deutsch im Kontakt und im Kontrast. Festschrift für Andrzej Kątny zum 65. Geburtstag*. Frankfurt a.M., 341–356.
- Brown, P./ S.C. Levinson (1987), *Politeness: Some universals in language use*. Cambridge.
- Carston, R. (2002), *Thoughts and utterances. The Pragmatics of explicit communication*. Oxford.
- Cheang, H.S./ M.D. Pell (2008), *The sound of sarcasm*. In: *Speech Communication* 50, 366–381
- Chen, A./ C. Gussenhoven/ T. Rietveld (2004), *Language-specificity in the perception of paralinguistic intonational meaning*. In: *Language and Speech* 4, 311–350.
- Culpeper, J. (1996), *Towards an anatomy of impoliteness*. In: *Journal of Pragmatics* 25 (3), 349–367.
- Culpeper, J. (2011), *It's not what you said, it's how you said it!: Prosody and impoliteness*. In: Linguistic Politeness Research Group (Hrsg.), *Discursive Approaches to Politeness*. Berlin etc., 57–83.
- Culpeper, J./ D. Bousfield/ W. Wichmann (2003), *Impoliteness revisited: with special reference to dynamic and prosodic aspects*. In: *Journal of Pragmatics* 35, 1545–1579.
- d'Avis, F./ J. Meibauer (2013), *Du Idiot! Din Idiot! Pseudo-vocative constructions and insults in German (and Swedish)*. In: B. Sonnenhauser/ H. Noel Aziz Hanna (Hrsg.), *Vocative!*, Berlin etc., 189–217.
- Goffman, E. (1986), *Interaktionsrituale. Über Verhalten in direkter Kommunikation*. Frankfurt a.M..
- Grauwunder, S./ B. Winter (2010), *Acoustic correlates of politeness: prosodic and voice quality measures in polite and informal speech of Korean and German speakers*. In:

- Proceedings of the International Conference for Speech Prosody 2010, Chicago 5/2010, Speech Prosody 2010, 100316: 1–4.
- Grice, M./ S. Baumann (2002), *Deutsche Intonation und GToBI*. In: Linguistische Berichte 191, 267–298.
- Grice, M./ S. Baumann/ R. Benz Müller (2005), *German Intonation in Autosegmental-Metrical Phonology*. In: S.-A. Jun (Hrsg.), *Prosodic Typology: The Phonology of Intonation and Phrasing*. Oxford, 55–83.
- Gussenhoven, C. (2014), *The Phonology of Tone and Intonation*. Cambridge.
- Gussenhoven, C. (2002), *Intonation and Interpretation: Phonetics and Phonology*. In: B. Bel/ I. Marlien. (Hrsg.). *Proceedings of Speech Prosody 2002*. Aix-En-Provence. 47–57.
- Haiman, J. (1998), *Talk is cheap: Sarcasm, alienation and the evolution of language*. New York.
- Haugh, M. (2015), *Im/Politeness Implicatures*. Berlin etc.
- House, J. (2008), *Constructing a Context with Intonation*. In: *Journal for Pragmatics* 38, 1542–1558.
- Kaufer, D. S. (1981), *Understanding ironic communication*. In: *Journal of pragmatics* 5, 495–510.
- Ladd, D. Robert (1996), *Intonational Phonology*. Cambridge.
- Leech, G. (1983), *Principles of Pragmatics*. Longman.
- Leggitt, J./ R. Gibbs (2000), *Emotional reactions to verbal irony*. In: *Discourse Processes* 29, 1–24.
- Mateo, J./ F. Yus (2013), *Towards a cross-cultural pragmatic taxonomy of insults*. In: *Journal of Language Aggression and Conflict* 1 (1), 86–113.
- McKinnon, S./ P. Prieto (2014), *The role of Prosody and Gesture in the Perception of Mock Impoliteness*. In: *Journal of Politeness Research* 10(2), 185–219.
- Morton, E.S. (1977), *On the occurrence and significance of motivation-structural rules in some bird and mammal sounds*. In: *The American Naturalist* 111, 855–869.
- Murray, I.R./ Arnott, J.L. (1993), *Toward the Simulation of Emotion in Synthetic Speech: A Review of the Literature on Human Vocal Emotion*. In: *Journal of the Acoustical Society of America*, 93(2), February, 1097–1108.
- Nakassis, C./ Snedeker, J. (2002), *Beyond sarcasm: Intonation and context as relational cues in children's recognition of irony*. In: A. Greenhill/ M. Hughs/ H. Littlefield/ H. Walsh (Hrsg.), *Proceedings of the Twenty-Sixth Boston University Conference on Language Development*. Somerville Mass, 429–440.
- Nowick-Dziewicka, E. (2012), *Banter and the Echo/Pretence Distinction*. In: E. Walaszewska/ A. Piskorska (Hrsg.), *Relevance Theory: More than understanding*. Cambridge, 245–259.
- Ohala, J.J. (1983), *Cross-language use of pitch: an ethological view*. In: *Phonetica* 40, 1–18.
- Ohala, J.J. (1984), *An ethological perspective on common cross-language utilization of F0 in voice*. In: *Phonetica* 41, 1–16.
- Ohala, J.J. (1994), *The frequency code underlines the sound symbolic use of voice of pitch*. In: L. Hinton et al. (Hrsg.), *Sound symbolism*, Cambridge. 325–247.
- Pierrehumbert, J./ M.E. Beckman (1988), *Japanese Tone Structure*. Cambridge, MA.

- Pierrehumbert, J. (1980), *The Phonology and Phonetics of English Intonation* (Ph.D thesis, MIT).
- Rockwell, P. (2000), *Lower, slower, louder: Vocal cues of sarcasm*. In: Journal of Psycholinguistics Research 29, 483–495.
- Rockwell, P. (2006), *Sarcasm and other mixed messages. The ambiguous ways people use language*. New York.
- Searle, J. R. (1982), *Indirekte Sprechakte*. In: J.R. Searle, Ausdruck und Bedeutung. Frankfurt a.M., 51–79.
- Sperber, D./ D. Wilson (1981), *Irony and the use-mention distinction*. In: P. Cole (Hrsg.) Radical Pragmatics. New York. Reprinted in S. Davis (Hrsg.) (1991) Pragmatics: A Reader. Oxford, 550–563.
- Sperber, D./ Wilson, D. (2002), *Relevance: Communication and Cognition*. Cambridge.
- Sperber, D. (1984), *Verbal irony: Pretense or echoic mention?* In: Journal of Experimental Psychology: General 113, 130–136.
- Stadler, S. A. (2006), *Multimodal (Im)politeness: The Verbal, Prosodic and Non-Verbal Realization of Disagreement in German and New Zealand English. Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy in Linguistics*. University of Auckland and Universität Hamburg.
- Wilson, D./ D. Sperber (1992), *On verbal irony*. In: Lingua 87, 53–76.
- Wilson, D./ T. Wharton (2006), *Relevance and Prosody*. In: Journal of Pragmatics 38, 1559–1578.
- Wilson, D. (2013), *Irony comprehension: A developmental perspective*. In: Journal of Pragmatics 59, 40–56.
- Žegarac, V. (1998), *What is “phatic communication”?* In: V. Rouchota/ A. Jücker (Hrsg.), Current Issues in Relevance Theory. Amsterdam etc., 327–362.
- Zwaan, R./ G. Radvansky (1998), *Situation models in language comprehension and memory*. In: Psychological Bulletin 123 (2), 162–185.