

AFFEKTÄUSSERUNGEN IN SPRACHKORPORA

Jürgen Trouvain

Universität des Saarlandes, Phonetik
trouvain@coli.uni-saarland.de

Abstract: Ziel der vorliegenden Studie ist eine Analyse von Affektäußerungen wie sie zum einen als gesondert aufgenommene Vokalisierungen in Spezialkorpora und zum anderen in natürlichen Dialog-Korpora vorkommen.

1 Einführung

Affektive Information in gesprochener Sprache erkennen wir als Hörer an bestimmten prosodischen Merkmalen wie mittlere Tonhöhe, Stimmumfang, Lautstärke und Timing sowie der Stimmqualität (vgl. [1]). Darüber hinaus gibt es aber vom sprachlichen Signal unabhängige Vokalisierungen wie Weinen, Stöhnen, Jauchzen etc., womit affektive Information übermittelt werden, die als "affect bursts" [2,3] oder Affektäußerungen bekannt sind. Sowohl für die Spracherkennung als auch für die Annotation von Sprachkorpora ist es von Interesse diese Art von Information korrekt zu interpretieren bzw. zu kodieren. Ebenso kann es für die Sprachsynthese relevant werden, solche Affektäußerungen in adäquater Weise zu verwenden (siehe auch [4]).

Am Beispiel von Lachen wird gezeigt, welche Unterschiede im akustischen Signal je nach Aufnahmesituation der betreffenden Datenbank feststellbar sind. Die von Schauspielern erzeugten Beispiele in der "Montréal Affective Voice Database" [5] zeigen sehr gut erkannte prototypische Varianten verschiedener Emotionsausdrücke. In der audio-visuellen Datenbank "AV Laughter Cycle" [6] (Anschauen lustiger Videoclips) kommt es zu einer größeren Variation der Lach-Vokalisierungen. Betrachtet man hingegen dialogische Sprache, in welcher die Interaktionspartner tatsächlich kommunizieren, indem sie beispielsweise eine gemeinsame Aufgabe zu lösen haben (z.B. [7,8]), so stellt man fest, dass sich Lachen a) in etlichen anderen akustischen Formen manifestiert als die Schauspieldaten vermuten lassen, und b) viel weniger intensive Formen beobachten lassen. Ein weiterer nicht zu unterschätzender Aspekt ist, dass es in Dialogen häufig zu Überlappungen kommt, von denen besonders Lachen betroffen ist. Aus diesem Grund ist es für eine korrekte Erfassung der Affektäußerung wichtig, dass bei Gesprächsdaten die Signale getrennt registriert und analysierbar gemacht werden (wie z.B. in [8]).

2 Affektivität in eigenständigen Äußerungen

Auch wenn die Übermittlung *textueller* Information im Mittelpunkt steht, so spielt der Transport affekt-bezogener Information durch das Sprachsignal sehr häufig eine wesentliche Rolle in der gesprochen-sprachlichen Kommunikation. Der Linguist Kenneth Pike hat diesen Punkt trefflich formuliert:

"[...] the hearer is frequently more interested in the speaker's attitude than in his words – that is, whether a sentence is 'spoken' with a smile' or with a sneer [...]" [9: p. 22]

Durch Verändern prosodischer Parameter wie Grundfrequenz, Timing, Intensität sowie der Stimmqualität werden gesprochene Äußerungen mit affektiver Information versehen. Neben der Artikulation verbalen Inhalts bringen wir in Spontansprache oftmals auch non-verbale Vokalisierungen hervor, die entweder einen kommunikativen Zweck verfolgen wie z.B.

Affektäußerungen oder wie vegetative Äußerungen. Interjektionen, die auch einen affektiven Charakter haben, stellen eine eigene verbale Kategorie dar. In Dialogen trifft man zudem häufig auf konventionalisierte Vokalisierungen wie "ja" oder "hm", die ebenfalls Affekt übertragen können.

2.1 Nicht-kommunikative Äußerungen

Ein Teil der non-verbalen Vokalisierungen umfasst solche phonetische Aktivitäten, die bei vegetativen Funktionen auftreten und wir daher nicht oder nur teilweise willentlich kontrollieren können. Diese Signale tragen zwar Informationen über den "Sprecher", aber sie dienen nicht der Kommunikation mit anderen. Zu diesen vegetativen Äußerungen gehören Husten, Schnarchen, Schlucken, Räuspern, Gähnen oder Niesen. Allerdings können Imitationen solcher Vokalisierungen zu kommunikativen Zwecken gezielt eingesetzt werden und damit auch zur Expressivität beitragen: absichtliche Gähn- oder Schnarchgeräusche zum Ausdruck von Langeweile oder ein Räuspern, um andere in diskreter Weise auf eine neue Situation aufmerksam zu machen.

Von diesen vegetativen Äußerungen zu trennen sind idiosynkratische Vokalisierungen, die ebenfalls nicht zu kommunikativen Zwecken eingesetzt werden. Als Beispiel mag hier ein inspiratorischer [s]-Laut dienen, wie sie beispielsweise Fußball-Bundestrainer Löw häufig in Interviews an Phrasengrenzen verwendet. Diese Vokalisierung ist aber von einem sehr ähnlich produzierten Laut zu unterscheiden, der auf Konventionen beruht und zum Ausdruck von "aua, das tut weh" benutzt wird.

2.2 Affektäußerungen

Ein großer Teil der non-verbalen Vokalisierungen dienen dem Ausdruck von Affekt. Scherer [2] nennt diese Vokalisierungen "affect bursts" und definiert sie als "very brief, discrete, nonverbal expressions of affect in both face and voice as triggered by clearly identifiable events".

Affektäußerungen sind teilweise als universal zu betrachten [10]. Lachen und Weinen kommt auch bei Menschen vor, die von Geburt an gehörlos sind. Dazu gehören weiterhin Jauchzen und Jubeln, aber auch Stöhnen und auch Schreien. All diese Arten von Vokalisierungen lassen sich leicht im Sport beobachten (vgl. [11, 12]).

Vokale Affektäußerungen der so genannten "Basisemotionen" Wut, Ekel, Angst, Freude, Traurigkeit und Überraschung sind über Kulturen hinweg gut identifizierbar und werden ähnlich gut erkannt wie deren visuelle Darbietungen [10, 13]. Neben der guten Erkennbarkeit gelten Affektäußerungen zusätzlich auch als sehr gut memorierbar [13]. Damit haben eigenständige Äußerungen mit affektivem Inhalt bei dessen Vermittlung einen Vorteil gegenüber artikulierter Sprache mit affektiver Prosodie, die nicht immer eindeutig erkannt wird.

2.3 Interjektionen

Merkmale der Wortklasse Interjektion sind Unflektierbarkeit, syntaktische Autonomie, Abwesenheit referentieller Bedeutung und der Umstand, dass sie zum Ausdruck einer spontanen Emotionen benutzt wird [14]. Typische Beispiele für Interjektionen sind "ach!", "naja!", "tja!", "ui!" oder "igitt!". Zusätzlich zu diesen primären Interjektionen gibt es auch sekundäre Interjektionen wie "meine Güte!" oder "Herrgott!", die als Wörter noch einen unabhängigen semantischen Wert besitzen.

Häufig weisen Interjektionen eine außergewöhnliche phonologische Struktur auf wie das vokallose "pst!" oder das mit Schnalzlauten produzierte "ts-ts". Interjektionen beruhen auf Konventionen, was auch das Beispiel "igitt" zeigt: ein Nicht-Muttersprachler des Deutschen

muss das Wort erlernen – im Gegensatz zu reinen Affektäußerungen, die universal verstanden werden.

Die Kurzäußerungen "oh" und "hm" werden auch als Interjektionen betrachtet. Hierbei ist es natürlich entscheidend, wie sich Dauer und Grundfrequenzverlauf verhalten, da sich dadurch die Bedeutung drastisch verändern kann (z.B. [15]). Auffallend ist der Befund, dass in Korpora beispielsweise "oh" meistens als Diskursmarker zum Signalisieren eines Informationsstatus eingesetzt wird und weniger zum Ausdruck von Emotionen [16]. Somit wird die prosodische Ausgestaltung solcher Kurzäußerungen in starkem Maße dazu eingesetzt, sowohl affektbezogene Information zu vermitteln als auch Funktionen in der Gesprächsinteraktion zu übernehmen [15,17]. Um den noch weitgehend ungeklärten Übergang zwischen Affektäußerungen und Interjektionen abzudecken ist der Begriff *Lautobjekte* vorgeschlagen worden [17].

2.4 Dialog-Signale

Etliche Kurzäußerungen dienen dazu, die Kommunikation zwischen zwei Gesprächspartnern aufrecht zu erhalten. Bestes Beispiel dafür sind Rückmeldesignale wie "ja" und "hm", die nicht immer eindeutig voneinander abgrenzbar als "backchannel utterances", "response tokens", "Rezipientensignale" oder "Hörervokalisierungen" bezeichnet werden. Diese Kurzäußerungen eignen sich in hervorragender Weise emotionale Nuancen und Einstellungen des Hörer-Sprechers kundzugeben [18,19].

3 Affektäußerungen in Korpora

3.1 Darstellen von Affekt durch Schauspieler

Eine gängige Methode zum Erfassen von Affektäußerungen besteht in Aufnahmen von Schauspielern [3, 5, 10], die bestimmte Emotionen ausdrücken sollen. Wie unterschiedlich die aufgenommenen Kategorien sein können zeigt Tabelle 1, in der drei verschiedene Studien mit Affektäußerungen [3, 5, 10] miteinander verglichen werden. Drei der sechs erwähnten "Basisemotionen" sind von allen Studien berücksichtigt, vier weitere Emotionen sind noch von zwei Studien beachtet (darunter die restlichen drei "Basisemotionen"). Es folgen weitere neun Emotionen, die jeweils nur von einer Studie als Kategorie als untersuchungswert eingestuft worden ist. Dieses Bild zeigt deutlich, dass sich bezüglich der wichtigsten Affekte keine Einigkeit besteht.

Zudem kann man davon ausgehen bzw. ist es erklärtes Ziel solcher Studien, dass es um das Erfassen *prototypischer* Varianten geht, die bestimmte affektive Zustände in voller Entfaltung ("full-blown" emotions) ausdrücken. Dieser Umstand ist wichtig, wenn man sich mit Formen affektiver Kurzäußerungen in realitätsnäheren Settings befasst.

Da im weiteren Verlauf des Aufsatzes vornehmlich auf Lachen, z.B. zum Ausdruck der Freude und Amüsiertheit, eingegangen wird, soll an dieser Stelle auf interessante Parallelen der wichtigen Emotionen "freudig" und "traurig" hingewiesen werden. Sowohl in der für die wissenschaftliche Welt frei zugänglichen "Montreal Affective Voice Database" [5] als auch bei Sauter und Kollegen [10] werden Lachen für "happy" und Weinen für "sad" verwendet. Neun der zehn Schauspieler der Montreal-Datenbank zeigen dabei eine auffällig ähnliche akustische Struktur für Lachen *und* für Weinen: ein Alternieren von "Vokal" und "Hauchlaut" sowie einer extrem hohen Grundfrequenz.

Schröder [3]	Belin et al. [5]	Sauter et al. [10]
Große Freude	happy	amusement
Wut	angry	anger
Ekel	disgusted	disgust
-	sad	sadness
-	surprised	surprise
-	fearful	fear
-	pleased	(sensual) pleasure
Erleichterung	-	relief
-	-	achievement/triumph
-	-	contentment
Bewunderung	-	-
Drohung	-	-
Langeweile	-	-
Schreck	-	-
Sorge	-	-
Verachtung	-	-
-	painful	-
-	(neutral)	-

Tabelle 1 – Emotionskategorien wie sie in drei Studien, die Affektäußerungen untersuchen, benannt wurden.

3.2 Anschauen lustiger Videoclips

Im Gegensatz zu Daten mit Schauspielern, die im Ausdruck von Affekt geübt sind, kann man weniger prototypische Daten auch durch Beobachten von Personen erlangen, die nicht im Ausdruck von Emotionen geschult sind. Verwendung findet hierbei das Anschauen von Filmmaterial, wodurch Versuchspersonen (Vpn) dazu stimuliert werden "gefühlter" Emotionen zu zeigen.

Bachorowski & Owren [20] ließen jeweils 2 Studenten gemeinsam amüsante Filmausschnitte (Dauer ca. 4 min) anschauen und haben dabei das Mikrofonsignal aufgenommen. Die Ergebnisse der 97 Vpn zeigen, dass mehrheitlich nicht stimmhaftes Lachen wie bei den Prototypen dominiert, sondern stimmlose Formen, die als "unvoiced snort-like" und "unvoiced grunt-like" bezeichnet werden. Die weiteren Analysen zeigen, dass die akustischen Ausformungen des Lachens eine *weitaus* größere Variabilität und Komplexität aufweisen als die Prototypen-Studien vermuten lassen.

Zu ähnlichen Ergebnissen kommt auch die Auswertung der "AV LaughterCycle database" einer Forschergruppe aus Mons [6], die einen ähnlichen Ansatz verfolgt wie [20] (diese Daten sind der wissenschaftlichen Welt frei zugänglich). Hier wurden 24 Personen gebeten sich (jeweils alleine) lustige Videoclips anzuschauen (ca. 12 Minuten Länge). Die Aufnahmen

bestehen zum einen aus dem Mikrofonsignal der Vpn und zum anderen aus visuell erfassten Motion Capture-Daten. Es handelt sich hierbei um halb-spontane Aufnahmen, da der Zweck allen Vpn vorher bekannt gewesen ist.

Auch wenn 34% aller 912 Lach-Ereignisse als "voiced" annotiert werden, so wird die Mehrzahl an Lachern einer von mehreren stimmlosen Kategorien wie "nasal", "breathy", "grunt-like" oder "quasi-silent" zugeordnet (45%).

Die beiden Studien mit dem Anschauen von Filmmaterial, das Lachen bei den Vpn erregen soll, zeigt sehr deutlich, dass es die *eine* Kategorie "Lachen" als solches gar nicht gibt, sondern dass Lachen als Affektäußerung in vielfältigen Formen zu Tage tritt. Interessanterweise kommt das prototypische stimmhafte Lachen nicht einmal mehrheitlich vor.

3.3 Dialogdaten

Als Beispiel für Daten mit Dialogen sollen zwei in Kiel aufgenommene und annotierte Datenbanken dienen. Zum einen handelt es sich um das "Kiel Corpus of Spontaneous Speech" [21], bei dem in einem Terminvereinbarungsszenario 16 Dialoge in einem "push-to-talk"-Modus (Überlappungen sind nicht möglich) aufgenommen wurden. Zum anderen handelt es sich um das "Lindenstraßen-Korpus" [IPDS], bei dem jeweils zwei Freunde in sechs Dialogen im Voll-Duplex-Verfahren (also mit möglichen Überlappungen) aufgenommen werden. Die Vpn hatten vor dem Gespräch die Aufgabe, Video-Schnipsel aus einer Episode der Seifenoper "Lindenstraße" in die richtige Reihenfolge zu bringen; im Dialog sollte man sich darüber unterhalten.

Bei beiden Datensammlungen ist es eindeutig, dass Lachen die dominierende, teilweise auch die einzige eigenständige Affektäußerung ist. Vergleicht man beide Korpora bezüglich der Anzahl des Lachens, so fällt auf, dass in dem viel größeren Kiel Corpus weniger Lachen gezeigt wird als in dem kleineren Lindenstraße-Korpus. Dieser Befund kann eventuell auf die Zusammensetzung der Dyaden zurückgeführt werden: die Sprecher im Lindenstraßen-Korpus waren befreundet.

Die Analyse der Lachereignisse im Kiel Corpus zeigt, dass fast ausschließlich Sprech-Lacher produziert werden, d.h. Lachen, das synchron mit Sprache produziert wird. Diese Sprech-Lacher sind im Lindenstraßen-Korpus nicht mehr dominant, aber mit 32% bzw. mit weiteren 13% Mischformen aus Sprech-Lachen und "freiem Lachen" noch stark vertreten. In Dialogdaten befinden sich also Formen von Lachen, die in den Nicht-Dialogen bzw. bei den Schauspielaffekten überhaupt nicht (oder nur sehr marginal) vorgekommen sind.

Lachen kann als Ausnahme des Dialog-Prinzips, dass nicht überlappend gesprochen wird, betrachtet werden. Im Kiel Corpus wurde diese Möglichkeit von vornherein ausgeschlossen – im Lindenstraße-Korpus hingegen sind 51% aller Lacher Ereignisse, bei denen beide Sprecher synchron lachen. Von den synchronen Lachern kommen 61% an Stellen vor, bei denen auch ein Sprecherwechsel stattfindet. Ein ähnliches Bild ergibt sich auch bei dem in [6] verwendeten Lego-Korpus. Es kann also vermutet werden, dass Lachen auch in der Funktion der Organisation von Redebeiträgen eine Rolle spielt.

Ein Blick auf Daten mit Dialogen zeigt also, dass die Vielfältigkeit der vokalen Ausdrucksform des Lachens noch reichhaltiger ist als die vorher betrachteten Korpora und Studien vermuten lassen. Zu dieser Erweiterung der Formmöglichkeiten muss auch die Variabilität bezüglich Intensität und Dauer gezählt werden, die sich in Dialogdaten stark erweitern kann. Dies lässt auch vermuten, dass eine beträchtliche Anzahl an Lachern nicht oder kaum zum Ausdruck von Affekt gebraucht wird, sondern eher gesprächssteuernde Funktionen übernimmt.

4 Zusammenfassung

Spezialkorpora und Dialogkorpora zeigen sehr deutliche Unterschiede in Bezug auf Vorkommen und Verwendung von Affektäußerungen. Etliche Affektkategorien kommen in vielen Korpora, die Dialoge enthalten, überhaupt nicht vor – beispielsweise zum Ausdruck von Traurigkeit, Angst oder Triumph. Das Beispiel des Ausdruckskomplexes "Lachen" zeigt, dass sich neben prototypischen Varianten noch etliche andere akustisch unterscheidbare Varianten in Spontan- und Halbspontansprache feststellen lassen. Zudem ist die prototypische Ausdrucksform des Lachens, wie sie in den Spezialkorpora vorkommt, nicht einmal die am häufigsten festzustellende Form in den Dialogen. In Gesprächen dominieren Varianten mit viel geringerer Intensität.

Ein weiterer wichtiger Aspekt ist, dass es sich zumindest beim Lachen nicht ausschließlich um eine Form der Expressivität des Sprechers handelt, die als Selbstkundgabe funktioniert. In nicht-überlappenden Dialogen kann man Lachen, meist als Sprech-Lachen geäußert, eine selbst-kommentierende Funktion zuweisen. In Voll-Duplex-Dialogen wird Lachen sehr häufig an Stellen von Sprecherüberlappungen zum Regulieren der sprachlichen Interaktion eingesetzt. Lachen ist somit also ein wichtiges Mittel der phatischen Funktion von Kommunikation. Die Funktionalität einer prototypischen Äußerungsform der Affektkategorien "glücklich" und "amüsiert" wird je nach Korpus dramatisch erweitert. Für eine korrekte Erfassung der Affektäußerung wichtig, dass bei Gesprächsdaten die Signale getrennt registriert und analysierbar gemacht werden (wie z.B. in [8]).

Starke Funktionserweiterung erfährt auch die eigentlich verbale Vokalisierung "ja", die neben ihrer lexikalischen Bedeutung der Zustimmung bzw. als Gradpartikel (z.B. in "ich hab's *ja* gewusst") auch zum Ausdruck von "Triumph" und "Dominanz" benutzt werden kann [12]. In diesen Funktionen findet man "ja" in den Dialogkorpora nicht, sondern "ja" wird dort als Diskursmarker eingesetzt: entweder einleitend beim Redebeitrag, weniger häufig den Redebeitrag beendend, und sehr häufig als Hörersignal. An dieser Stelle sei auf eine erfolgreiche prosodische Modellierung von "ja" hingewiesen, die auch für Synthesezwecke genutzt werden kann [22].

Was lässt sich weiterhin für die Sprachtechnologie, z.B. für Dialogsysteme, ableiten? Auch wenn sich die Affektäußerungen in Spezialkorpora und in Dialogkorpora sehr stark voneinander unterscheiden, so ist doch die Betrachtung der Korpus-Arten wichtig. Mit Sicherheit lassen sich Affektäußerungen als gut erkennbare und isolierte "cue phrases" einsetzen [23, 24]. Der Einsatz von Affektäußerungen in direkter Nachbarschaft zu Sprachäußerungen verlangt hingegen eine phonetische Anpassung an die verbalen Abschnitte [19, 23]. Bezüglich der Synthetisierung von natürlich klingendem Lachen sind die Ergebnisse immer noch bescheiden, was aktuelle Tests mit verschiedenen Methoden zeigen [25].

Literatur

- [1] Cowie, R., Douglas-Cowie, E. & Schröder, M. (eds.) (2000). Proceedings of the ISCA Workshop on Speech and Emotion: A Conceptual Framework for Research, Northern Ireland 2000. Belfast, NI: Textflow.
- [2] Scherer, K. R. (1994). Affect Bursts. In: van Goozen, S. H. M., van de Poll, N. E. & Sergeant J. A. (Eds.). *Emotions*, Lawrence Erlbaum, Hillsdale, NJ, pp. 161-193.
- [3] Schröder, M. (2003) Experimental study of affect bursts. *Speech Communication* 40(1-2), pp. 99-116.
- [4] Burkhardt, F. & Stegmann, J. (2009) Emotional speech synthesis: Applications, history and possible future. *Tagungsbände der 20. ESSV (Band 1)*, Dresden, 190-199.

- [5] Belin P., Fillion-Bilodeau S., Gosselin F. (2008) The Montreal Affective Voices: A validated set of nonverbal affect bursts for research on auditory affective processing Behavior Research Methods 40(2), pp 531-539.
- [6] Urbain, J. et al. (2010) The AVLaughterCycle Database. Proc. 7th international Conference on Language Resources and Evaluation (LREC), Malta, May 2010.
- [7] Kehrein, R. (2002) Prosodie und Emotionen. Tübingen: Niemeyer.
- [8] IPDS (2006) Video Task Scenario: Lindenstrasse – The Kiel Corpus of Spontaneous Speech, Volume 4, DVD, Institut für Phonetik und Digitale Sprachsignalverarbeitung Universität Kiel.
- [9] Pike, K. (1945). The Intonation of American English. Ann Arbor: University of Michigan Press.
- [10] Sauter, D.A., Eisner, F., Ekman, P. & Scott, S.K. (2010). Cross-cultural recognition of basic emotions through nonverbal emotional vocalizations. PNAS – Proceedings of the National Academy of Sciences, 107(6), pp. 2408-2412.
- [11] Braun, A. 2001. Phonetische Betrachtungen zu einem Phänomen im Tennissport: Eine explorative Studie zum grunting. In: Mauelshagen, C. & Seifert, J. (ed.): Sprache und Text in Theorie und Empirie. Beiträge zur germanistischen Sprachwissenschaft. Stuttgart: Steiner, pp. 198-208.
- [12] Trouvain, J. (in Vorb.) Phonetische Betrachtungen von Torjubel bei WM-Kommentatoren.
- [13] Armony, J.L., Chochol, C., Fecteau, S. & belin, P. (2007). Laugh (or cry) and you will be remembered. Psychological Science 18 (12), pp. 1027-1029.
- [14] Nübling, D. (2004). Die prototypische Interjektion: Ein Definitionsvorschlag. Zeitschrift für Semiotik 26, pp. 11-45.
- [15] Ward, N. (2006). Non-lexical conversational sounds in American English. Pragmatics & Cognition 14, pp. 129-182.
- [16] Norrick, N. (2008). Interjections as pragmatic markers. Journal of Pragmatics 41(5), pp. 866-891.
- [17] Reber, E. & Couper-Kuhlen, E. (2009). Interjektionen zwischen Lexikon und Vokalität: Lexem oder Lautobjekt? In: A. Deppermann & A. Linke (Hgg.), *Sprache intermedial: Stimme und Schrift, Bild und Ton*. Berlin & New York: de Gruyter, pp. 69-96.
- [18] Schröder, M., Heylen, D. & Poggi, I. (2006). Perception of non-verbal emotional listener feedback. Proc. Speech Prosody 2006, Dresden.
- [19] Pammi, S. & Schröder, M. (2009). Annotating meaning of listener vocalizations for speech synthesis. Proc. International Conference on Affective Computing and Intelligent Interaction, Amsterdam.
- [20] Bachorowski, J.-A., Smoski, M.J. & Owren, M.J. (2001). The acoustic features of human laughter. Journal of the Acoustical Society of America 111(3), pp.1582-1597.
- [21] IPDS (1995-1997). The Kiel Corpus of Spontaneous Speech, Volume 1-3, CDRom#2-4.
- [22] Stocksmeier, T., Kopp, S. & Gibbon, D. (2007). Synthesis of Prosodic Attitudinal Variants in German Backchannel *ja*. Proc. Interspeech Antwerpen, pp. 1290-1293.
- [23] Trouvain, J. & Schröder, M. (2004). How (not) to add laughter to synthetic speech. Proceedings of the Workshop on Affective Dialogue Systems, Kloster Irsee, 229-232.
- [24] Campbell, N. (2006). Conversational speech synthesis and the need for some laughter. IEEE Transactions on Audio, Speech, and Language Processing 14 (4), pp. 1171-1177.
- [25] Cox, Trevor (2010). Laughter synthesis. New Scientist. <http://www.soundsfunny.org/turing/results.php>