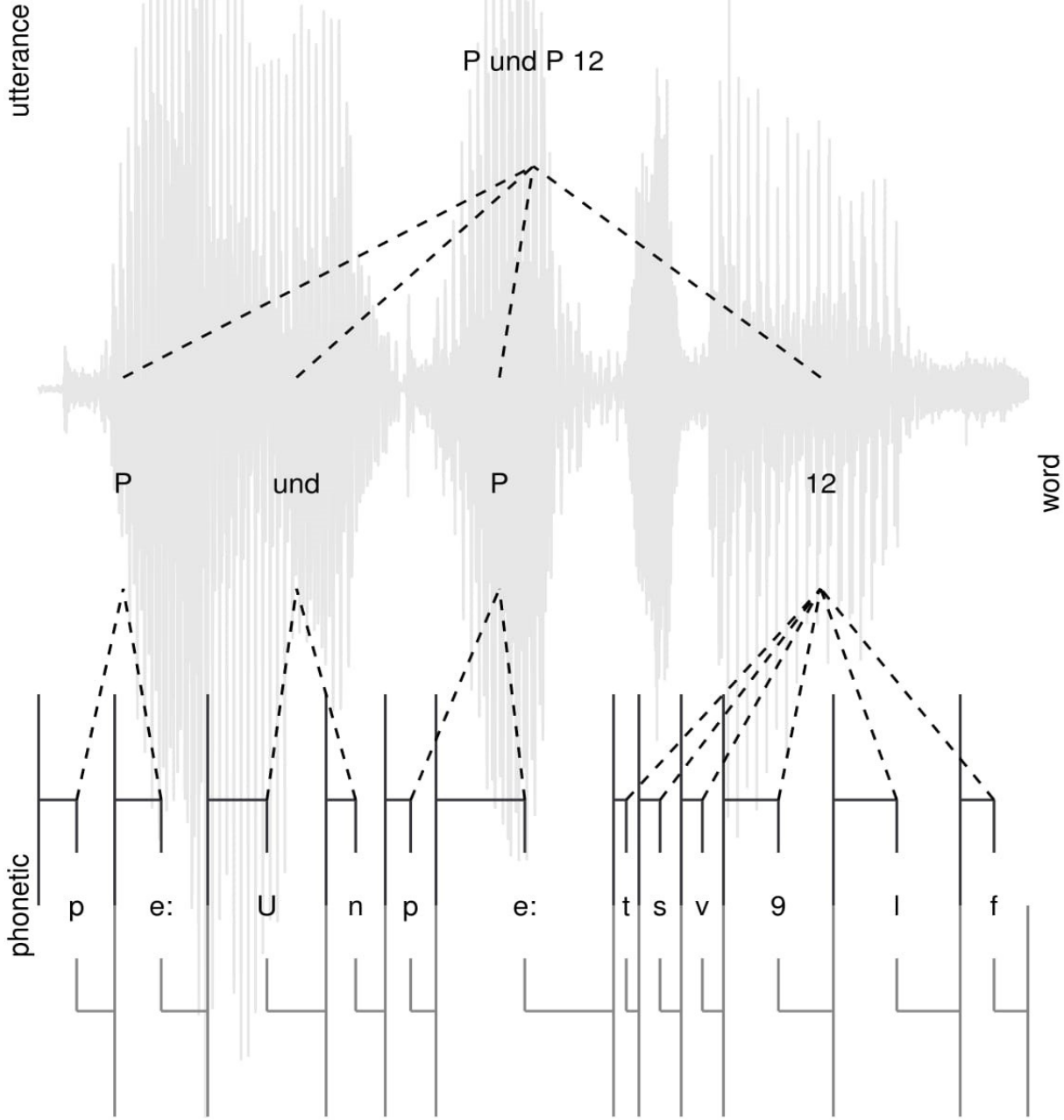
Book of Abstracts

12. Tagung Phonetik und Phonologie im deutschsprachigen Raum



12. - 14. Oktober 2016

München, Deutschland

Publiziert von:
Institut für Phonetik und Sprachverarbeitung
Ludwig-Maximilians-Universität München

Tagungsinformationen:
http://www.phonetik.uni-muenchen.de/institut/veranstaltungen/pundp12/

Organisationsteam
Christoph Draxler, Markus Jochim, Felicitas Kleber

Wissenschaftliches Komitee
Christoph Draxler
Nikola Eger
Markus Jochim
Felicitas Kleber
Sandra Peters
Raphael Winkelmann

Redaktion des Book of Abstract
Katharina Schmidt

VORWORT

[sɛɐvʊs gʀɪas di]

Wir freuen uns sehr Sie in diesem Jahr als Gastgeber der 12. Tagung „Phonetik und Phonologie im deutschsprachigen Raum“ an der LMU München und dem Institut für Phonetik und Sprachverarbeitung begrüßen zu dürfen.

Die P&P hat sich seit dem ersten Treffen 2004 in Potsdam als wichtige Plattform für ForscherInnen und NachwuchswissenschaftlerInnen etabliert, die vor allem in Deutschland, Österreich und der Schweiz an der Schnittstelle zwischen Linguistik und Phonetik oder aber direkt an Fragen zur Phonetik und Phonologie des Deutschen und seiner regionalen Varietäten und Dialekte arbeiten.

Neben Beiträgen aus allen Bereichen der Phonetik und Phonologie wird auf der diesjährigen Tagung der thematische Schwerpunkt "Datenbanken, Korpora und Big Data" mit einer eigenen Session aufgegriffen. Mit diesem Motto sollen gezielt auch technische und informationsverarbeitende Fachbereiche angesprochen (Stichwort *Digital Humanities*) und der Bedeutung von Empirie in Phonetik und Phonologie Rechnung getragen werden. Darüberhinaus spiegelt der Fokus auf zunehmend größere Datenbanken und empirische Methoden die Lehr- und Forschungsausrichtung des Instituts für Phonetik und Sprachverarbeitung der LMU München als auch die aktuellen Forschungsansätze in der Wissenschaft wider.

Das vorliegende *Book of Abstracts* enthält neben allgemeinen Informationen zu Tagungsprogramm und Tagungsort alle Abstracts zu den auf der Tagung präsentierten Vorträgen und Postern. Zu vielen Abstracts gibt es zusätzlich Konferenzartikel, die im Tagungsband zur diesjährigen P&P12 erschienen sind – auf diese wird unter den jeweiligen Abstracts hingewiesen.

Wir bedanken uns herzlich bei allen AutorInnen für die zahlreich eingegangenen Beiträge, freuen uns über die große Teilnehmerzahl und wünschen uns allen eine interessante und anregende Tagung.

Felicitas Kleber und Christoph Draxler
München, Oktober 2016

GRUSSWORT

It is a great pleasure to welcome delegates to this P&P Conference in Munich. It is also very pleasing to see how this conference series has flourished since the first P&P in Potsdam in 2004.

This visionary idea of holding these yearly conferences has now culminated in the 12th P&P in Munich with 3 keynote papers, oral presentations spanning two days, as well as three workshops with over 50 poster presentations across so many areas: human speech processing, speech corpora and tools, models of speech production and perception, prosody, first and second language acquisition - indeed many of the major sections that we typically find in the international congress of phonetic sciences.

Perhaps most importantly, the P&P - in a way that is similar to the regular Acoustical Society of America meetings - provides such a valuable opportunity to present research that is still in progress and to discuss and to develop ideas with the help of colleagues that now attend these conferences from such a range of different disciplinary perspectives.

The great progress that has been made in P&P over more than a decade provides a very firm foundation for extending its international reach in the years to come - thereby reflecting the very high standards of scientific research in phonetics and phonology that is carried out in Germany.

I wish everyone a productive and enjoyable time here in Munich.

Jonathan Harrington
Chair of Phonetics and Speech Processing
LMU Munich

SPONSOREN

amazon

münchener universitätsgesellschaft

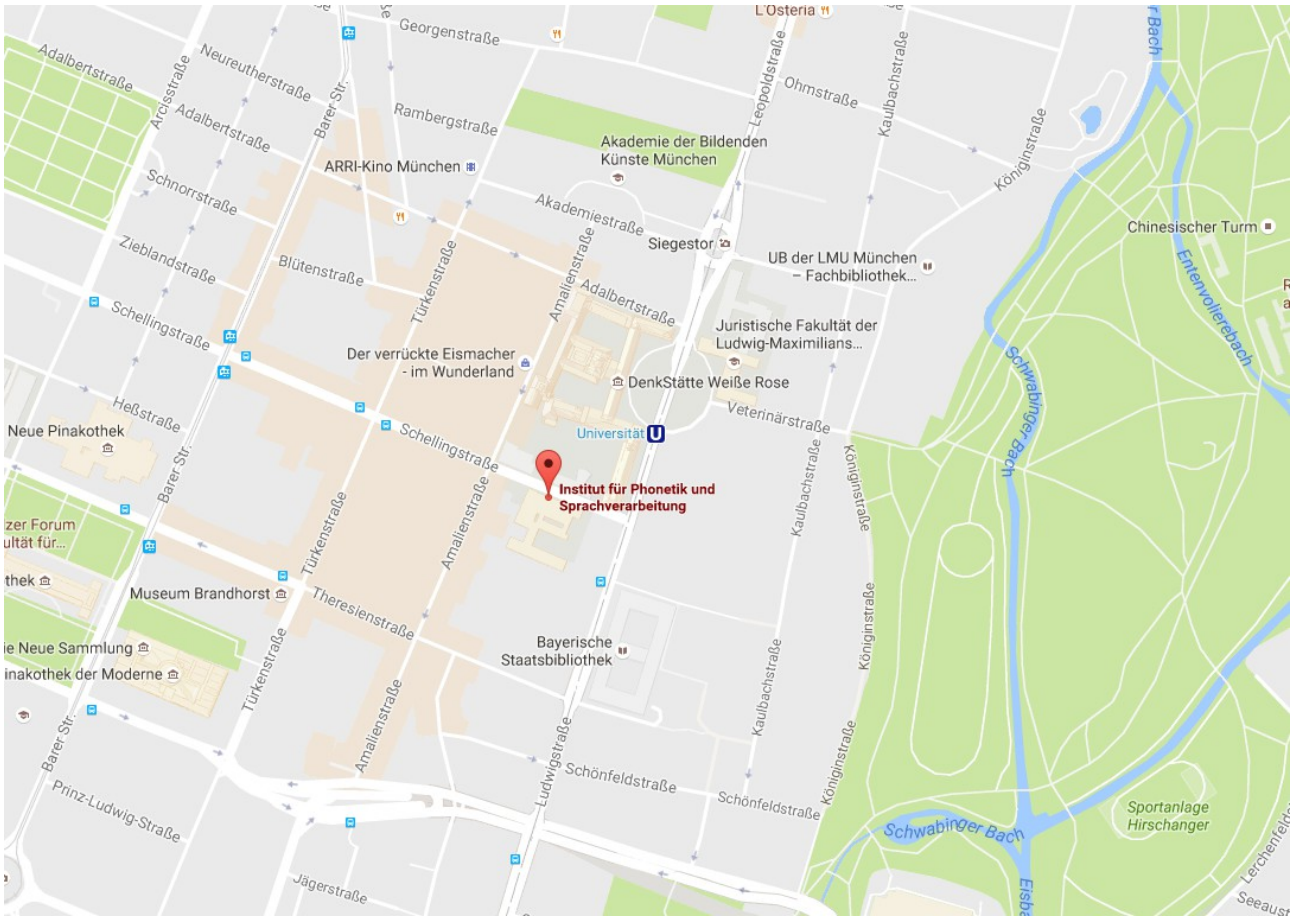
Gesellschaft von Freunden und Förderern der Ludwig-Maximilians-Universität München e.V.

IPS

INSTITUT FÜR PHONETIK UND SPRACHVERARBEITUNG

LMU LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

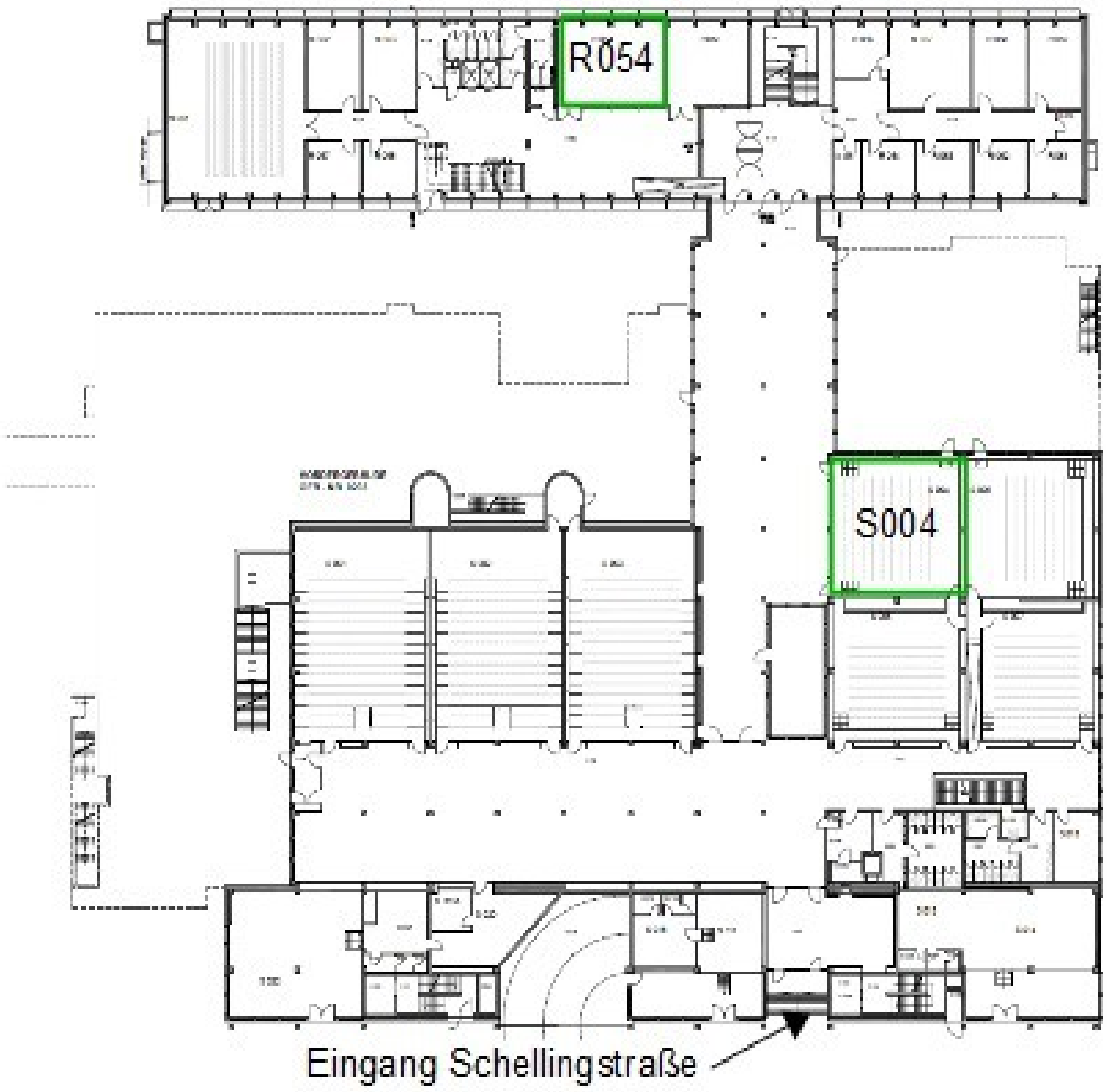
TAGUNGSORT



Ludwig-Maximilians-Universität München
Institut für Phonetik und Sprachverarbeitung
Schellingstr. 3
80799 München

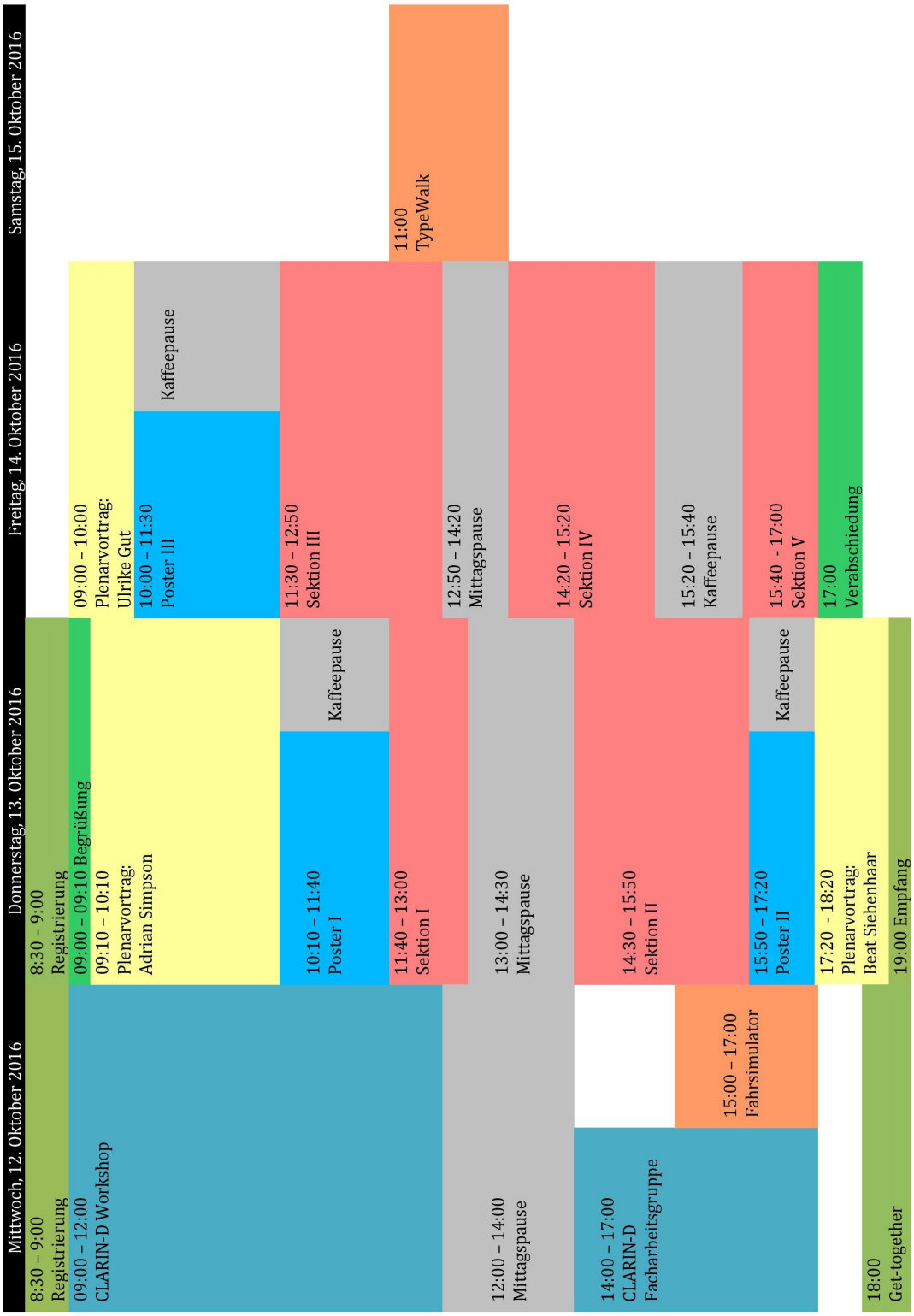
Anfahrt vom Hauptbahnhof:
Folgen Sie dem grün unterlegten S-Zeichen zur S-Bahn. Am Automaten erhalten Sie eine Einzelfahrkarte (1 Zone), eine Streifenkarte (wenn Sie mehrere Fahrten planen) oder eine Tageskarte (für Gruppen empfohlen). Auf der Streifenkarte sind pro Person zwei Streifen zu entwerten. Nehmen Sie eine S-Bahn Richtung Innenstadt/Marienplatz. Am Marienplatz steigen Sie in die U-Bahn Richtung Münchner Freiheit (U3, U6) um. Fahren Sie zwei Stationen bis Universität und wählen Sie den Ausgang zur Schellingstraße. Biegen Sie nach rechts in die Schellingstraße ein. Das Institut ist ca. 50 m weiter auf der linken Seite. Der Eingang (Hausnummer 3) befindet sich bei den Treppenstufen. Nach dem Betreten des Gebäudes gehen Sie durch die vor Ihnen liegende Glastüre. Nun sind sie im Areal der P&P angekommen.

RÄUMLICHKEITEN



- **R054:** Seminarraum für CLARIN-D-Veranstaltungen (Mittwoch) und Tagungsbüro (Donnerstag, Freitag)
- **S004** Vortragssaal
- Die **Posterpräsentationen** finden im lange Gang vor dem Vortragssaal statt.

TAGUNGSÜBERSICHT



TAGUNGSPROGRAMM

Tag 1: Donnerstag, 13. Oktober 2016

08:30 - 09:00 **Registrierung**

09:00 - 09:10 **Begrüßung**

Felicitas Kleber, Christoph Draxler

09:10 -10:10 **Plenarvortrag I**

Adrian Simpson
Describing and accounting for sound patterns in spontaneous speech

10:10 - 11:40 **Poster I und Kaffeepause**

- *Melanie Weirich, Adrian Simpson*
  Changes in IDS and ADS during parental leave – project sketch and first results of pilot studies
- *Kim Strütjen, Ruben van de Vijver*
  Vowel confusions in noise by German listeners: A study of oral and nasalized vowels
- *Gerrit Kentner, Isabelle Franz, Christian Dück*
  Der optionale Komplementierer im Deutschen - ein Fall prosodischer Syntax
- *Louise Probst, Angelika Braun*
  Geflüsterte Angst und behauchte Trauer – Stimmqualität und Emotionen
- *Eugen Klein, Jana Brunner, Phil Hoole*
  Relation between articulatory and acoustic information in phonemic representations
- *Johanna Stahnke*
  Prosodic variation in conceptual distance and proximity: Self-repairs in French
- *Frederike Urke, Henning Reetz, Gea De Jong-Lendle*
  Die Wahrnehmung reduzierter Sprache unter Rauschen
- *Laura Sichlinger*
  Untersuchung des Kompensationsverhaltens bei Echtzeitmanipulation der Zeitstruktur des auditorischen Feedbacks
- *Uwe Reichel, Jennifer Cole*
  Entrainment analysis of categorical intonation representations
- *Katharina Nimz, Judith Baumann, Arkadiusz Rojczyk*
  Universal phonetics revisited: Eine cross-linguistische Untersuchung zum Einfluss der Stimmhaftigkeit des Folgekonsonanten auf die Vokallänge im Polnischen und Deutschen

- *Susanne Beinrucker*
  Schwa Elision in German Utterances of Bilingual Speakers with Different Ambient Languages during Speech Acquisition
- *Michael Fischbacher, Eva Reinisch*
  The temporal relation between talker and word recognition
- *Miriam Oschkinat*
  Musikalität und Auditory Acuity
- *Bettina Hobel, Sylvia Moosmüller*
  The realisation of Albanian laterals in German as a second language: A case study
- *Conceição Cunha, Nicole Weidinger, Herta Alejo, Katharina Hogrefe, Wolfram Ziegler, Katrin Lindner, Georg Goldenberg*
  The development of intonation and narrative abilities in a narrative retelling corpus
- *Stephan Schmid*
  Wie Deutschschweizer Lernende die stimmhaften Obstruenten des Italienischen aussprechen
- *Marie-Christin Himmel, Barış Kabak*
  Categorial and acoustic changes to rhoticity in bilinguals
- *Katalin Mády, Uwe D. Reichel*
  How to distinguish between self- and other-directed wh-questions?
- *Christine Mooshammer, Tamara Rathcke*
  Opa vs Oper: Neutralization of /ɐ/ and unstressed /a/ contrast in a perception and production study?

11:40 - 13:00 **Sektion I**

11:40 - 12:00 *Zofia Malisz, Frank Zimmerer, Emmanuel Ferragne, François Pellegrino, Erika Brandt, Bistra Andreeva, Bernd Möbius*
Ist gleich schnell gleich schnell? Perzeption intendierter und realisierter Sprechrate von deutschen und französischen Sprechern

12:00 - 12:20 *Nikola Anna Eger, Eva Reinisch*
German learners' productions of English sound contrasts: The role of acoustic properties on accent ratings

12:20 - 12:40 *Christina Domene Moreno, Baris Kabak*
L3 phonology - cross-language, psychoacoustic,, extralinguistic factors

12:40 - 13:00 *Frank Zimmerer, Jeanin Jügler*
Ein computer-gestützter Vokaltrainer für die deutsche Sprache

---

13:00 - 14:30 **Mittagessen**

14:30 - 15:50 **Sektion II**

14:30 - 14:50 *Caroline Kaufhold, Christine Martindale, Axel Horndasch, Klaus Reinhard, Elmar Nöth*
PATSY-I: A Corpus on Non-Native English Air Traffic Communication

14:50 - 15:10 *Laura Fernández Gallardo*
Recording a High-Quality German Speech Database for the Study of Speaker Personality and Likability

15:10 - 15:30 *Jürgen Trouvain, Frank Zimmerer*
Phonetic learner corpora for research on German as a foreign language

15:30 - 15:50 *Adrian Leemann, Marie-José Kolly*
Big Data for analyses of small-scale regional variation: A case study on sound change in Swiss German

---

15:50 - 17:20 **Poster II und Kaffeepause**

- *Petra Wagner, Aleksandra Ćwiek, Barbara Samlowski*
  Beat it! – Gesture-based Prominence Annotation as a Window to Individual Prosody Perception
- *Grigorij Aronov, Antje Schweitzer*
  Acoustic correlates of word stress in German spontaneous speech
- *Katrin Wolfswinkler, Eva Reinisch*
  The impact of accent familiarity on the perception of difficult sound contrasts for German learners of English
- *Sebastian Bredemann*
  A tonal analysis of the Limburgian Dialect spoken in Reuver
- *Jan Michalsky*
  Perception of Pitch Scaling in Rising Intonation. On the Relevance of f0 Median and Speaking Rate in German
- *Riccarda Funk, Christina Otto*
  Die akustischen und artikulatorischen Korrelate des /r/ im Norddeutschen. Eine Ultraschallstudie.
- *Michael Pucher, Michaela Rausch-Supola, Sylvia Moosmüller, Markus Toman, Dietmar Schabus, Friedrich Neubarth*
  Open data for speech synthesis of Austrian German language varieties
- *Denis Arnold, Fabian Tomaschek*
  The Karl Eberhards Corpus of spontaneously spoken southern German in dialogues – audio and articulatory recordings
- *Nina Poerner, Florian Schiel*
  An automatic chunk segmentation tool for long transcribed speech recordings

- *Elissa Pustka, Christoph Gabriel, Trudel Meisenburg*
  Romance Corpus Phonology: from (Inter-)Phonologie du Français Contemporain (I)PFC to (Inter-)Fonología del Español Contemporáneo (I)FEC
- *Iona Gessinger, Eran Raveh, Johannah O'Mahony, Ingmar Steiner, Bernd Möbius*
  A Shadowing Experiment with Natural and Synthetic Stimuli
- *Yshai Kalmanovitch*
  The Zurich Tangram Corpus (ZTC): Speech in interaction and phonetic convergence
- *Johanna Dobbriner, Oliver Jokisch, Michael Maruschke*
  Assessment of Prosodic Attributes in Codec-Compressed Speech
- *Daniel Duran, Natalie Lewandowski, Antje Schweitzer*
  Wahrnehmungsexperimente mit Hilfe eines Computerspiels
- *Urban Zihlmann, Adrian Leemann*
  'Chend' met – 'Kind' mit : using Big Data to explore phoneme-to-grapheme mapping in Lucerne Swiss German
- *Markus Jochim, Christoph Draxler*
  Fully Automated Accent Correction for Computer-Assisted Speech Rhythm Training: A Pilot Study
- *Volker Dellwo*
  PresenterPro: A Praat plug-in for efficiently presenting and recording speech prompts
- *Mathias Walther, Jokisch Oliver, Taieb Mellouli*
  Two-stage Decision Trees for Automatic Speaker Likability Classification
- *Hiroyuki Tanaka, Tamara Rathcke*
  Then, what is charisma? The role of audio-visual prosody in L1 and L2 political speeches

---

17:20 – 18:20 **Plenarvortrag II**

Beat Siebenhaar

Automatisches Alignment – eine projektspezifische Usersicht

19:00 **Empfang**

Tag 2: Freitag, 14. Oktober 2016

09:00 – 10:00 **Plenarvortrag III**

Ulrike Gut

Corpus Phonology - using corpora in phonological and phonetic research

10:00 – 11:30 **Poster III und Kaffeepause**

- *Oxana Rasskazova, Malte Belz, Christine Mooshammer, Jelena Krivokapić*
  Acoustic and articulatory manifestations of final lengthening and voicing contrasts for German learners of English as a second language
- *Tamara Rathcke, Florent Chevalier, Jane Stuart-Smith*
  What is the fate of the Scottish Vowel Length Rule in Glasgow?
- *Tabea Thies, Anne Hermes, Doris Mücke*
  Coordination Defictis in Essential Tremor Patients with Deep Brain Stimulation
- *Gerrit Kentner*
  New evidence for prosodic parallelism affecting German morphophonology
- *Benjamin Weiss*
  Voice Descriptions by Non-Experts: Validation of a Questionnaire
- *Aleksandra Cwiek, Sina Neueder, Petra Wagner*
  Investigating the communicative function of breathing and non-breathing “silent” pauses
- *Cornelia J. Heyde, James M. Scobbie*
  Wenn Stotterer nicht stottern. Quantifizierung dynamischer Ultraschalldaten
- *Renate Raffelsiefen, Anja Geumann*
  AI vs. AU in American English compared to German
- *Simon Betz, Petra Wagner, Jana Voße*
  Deriving a strategy for synthesizing lengthening disfluencies based on spontaneous conversational speech data
- *Katharina Nimz, Kai Ole Koop, Katharina Immel*
  Wer die Qual hat, hat keinen Wal: Orthographische Effekte bei der Produktion deutscher Vokale
- *Jan Michalsky, Heike Schoormann*
  Effects of perceived attractiveness and likability on global aspects of fundamental frequency
- *Amra Odobasic*
  Vocal Fry: A Marker of Sophistication or Stupidity?
- *Nicola Klingler*
  Der Einfluss der F0-Kontur als akustischer cue auf die Perzeption chinesischer Deutschlerner in konkurrierenden Kontexten

- *Carolin Schmid*
  German initial laterals by bilingual L1 Bosnian migrants in Vienna
- *Katalin Mády, Felicitas Kleber, Uwe Reichel, Ádám Szalontai*
  The interplay of prominence and boundary strength: a comparative study

---

11:30 – 12:50 **Sektion III**

11:30 - 11:50 *Bettina Braun, Yuki Asano, Nicole Dehé*
Lexical and prosodic operators to contrastive alternatives

11:50 - 12:10 *Isabelle Franz, Gerrit Kentner, Frank Domahs*
The impact of animacy and rhythm on the linear order of conjuncts in child language.

12:10 - 12:30 *Elina Rubertus, Dzhuma Abakarova, Jan Ries, Aude Noiray*
The development of coarticulation in German children

12:30 - 12:50 *Susanne Fuchs, Uwe D. Reichel*
On the relationship between pointing gestures and speech production in German counting out rhymes: Evidence from motion capture data and speech acoustics

---

12:50 - 14:20 **Mittagessen**

14:20 - 15:20 **Sektion IV**

14:20 - 14:40 *Hannah Leykum, Sylvia Moosmüller*
(Mor-)phonotactic consonant clusters in Standard Austrian German and Standard German German

14:40 - 15:00 *Stefanie Jannedy, Melanie Weirich*
The Acoustics of Northern German Fricative Contrasts

15:00 - 15:20 *Heike Schoormann, Wilbert Heeringa, Joerg Peters*
Monolingual and trilingual production of Northern Standard German vowels

---

15:20 - 15:40 **Kaffeepause**

15:40 - 17:00 **Sektion V**

15:40 - 16:00 *Benno Peters, Matthias Hoffmann, Laura-Marie Andresen*
Sprachdatenerhebung und Kontextvariation: Frageintonation in den Kontexten Dominanz und Unterordnung

16:00 - 16:20 *Christine T. Röhr, Tabea Thies, Stefan Baumann, Martine Grice*
Prosodic Marking of Information Status in Task-Oriented Dialogues

16:20 - 16:40 *Petra Wagner, Katalin Mády, Ádám Szalontai*
Teasing apart lexical stress and sentence accent in Hungarian and German

16:40 - 17:00 *Frank Kügler, Stefan Baumann, Antje Schweitzer, Petra Wagner*
DIMA – Erweiterte Richtlinien zur Annotation deutscher Intonation

EINGELADENE VORTRÄGE

Describing and accounting for sound patterns in spontaneous speech

Adrian Simpson

Institute of German Linguistics, Friedrich Schiller University Jena, Germany
Adrian.Simpson@uni-jena.de

This talk explores the range of sound patterns found in spontaneous speech, concentrating primarily on the variation in segmental aspects of words and phrases. These patterns have been found to correlate with a number of different factors, e.g. frequency of occurrence, place in prosodic structure, pragmatic function. Many of these segmental patterns are treated as resulting from reductions in the size, number or temporal extent of articulatory gestures associated with a lexical item. Reductions, in turn, have been seen as the products of ease of articulation, with the speaker orienting more towards the system-oriented end of Lindblom's (1990) H&H continuum. However, different phonetic realisations of the same word may represent more than just a speaker exploiting a range of different reduction possibilities and may instead be due to more complex long-domain patterns. This will be illustrated using examples of co-occurrence restrictions on glottal stops from spontaneous Suffolk English (Simpson 2006). A further problem with ease of articulation is that the result of a reduction can also result in an articulation that is apparently more complex than the unreduced articulation. One widespread and commonly reported pattern illustrating this is the reduction of a stop to a fricative which would seem to require more articulatory effort than the unreduced stop. The systematic intervocalic fricative realisation of /t/ in SSBE (e.g. *British, political*) is a good contemporary example (Simpson 2007). In connection with this, we will ask why more fricatives aren't in fact reduced to stops.
Finally, we will examine what the boundaries on reduction are. So, for example, recent studies have described reductions of filler words, such as *eigentlich* in German (Kohler 2001) or *eigenlijk* in Dutch (Plug 2006, Ernestus, in press), showing that a polysyllabic word can be reduced to little more than a single syllable. Furthermore, such forms are often incomprehensible if they are played to listeners in isolation, extracted from the utterance context. Despite the presence of such extreme reductions, it is not clear what the restrictions on reduction in general are, i.e. what constitutes wellformed reduced speech. Pathological speech data from a speaker in the early stages of atypical Parkinson's disease show that incorrect reduction patterns can have serious adverse effects on comprehension.

References

[1] Ernestus, Miriam. In press. Qualitative and quantitative aspects of phonetic variation in Dutch *eigenlijk.* In F.
[2] Cangemi, M. Clayards, O. Niebuhr, B. Schuppler & M. Zellers (eds.) *Rethinking reduction.* Berlin/Boston: De Gruyter Mouton.
[3] Lindblom, Björn. 1990. Explaining phonetic variation: a sketch of the H and H theory. In William J. Hardcastle & Alain Marchal (eds.), *Speech production and speech modeling.* Dordrecht: Kluwer Academic Publishers, 403–439.
[4] Kohler, Klaus J., 2001. Articulatory dynamics of vowels and consonants in speech communication. *Journal of the International Phonetic Association*, *31*(01), 1–16.
[5] Plug, Leendert. 2006. From words to actions: the phonetics of *eigenlijk* in two communicative contexts. *Phonetica*, *62*(2-4), 131–145.
[6] Simpson, A.P., 2006. Phonetic processes in discourse. In K. Brown (ed.) *Encyclopedia of Language and Linguistics 2 ed.,* vol. *9,* Amsterdam: Elsevier, 379-385.
[7] Simpson, Adrian P. 2007. Phonetische Motivation für Lenition: Gebrauch oder Missbrauch von phonetischen Erklärungen? In P. Gallmann, C. Lehmann, & R. Lühr (eds.), Sprachliche Motivation. *Zur Interdependenz von Inhalt und Ausdruck.* Tübinger Beiträge zur Linguistik, Band 502, Tübingen: Narr, 211–223.

Automatisches Alignment – eine projektspezifische Usersicht

Beat Siebenhaar

Institut für Germanistik, Universität Leipzig
siebenhaar@uni-leipzig.de

Seit Jahren, Jahrzehnten, hoffen Phonetiker_innen, dass ihnen die digitalen Helfer die mühselige und fehleranfälligen Arbeitsschritte der Transkription, der Segmentation und des Etikettierens abnehmen. Die Fortschritte der digitalen Signalverarbeitung bringen uns diesem Ziel immer wieder ein bisschen näher. – Vielen Dank hier an das Münchner Team, das uns MAUS u.a. zugänglich macht. Die Hoffnungen sind aber wohl immer größer, als das, was aktuell möglich ist. Ich will aus der Sicht eines hoffnungsvoll enttäuschten Hoffenden zeigen, was ich mit einem ganz konkreten Projekt zu 'Sprechtempo und Reduktion im Deutschen (SpuRD)' im Hintergrund erhoffe, welche Erwartungen ich einem automatischen Alignment entgegenbringe und welches die (meine) Realität ist. Mit meinen Mitarbeiter_innen nutze ich MAUS intensiv; all unsere nun über 150.000 Segmente sind aber manuell aufgearbeitet. Ich zeige mit diesem Vortrag aus der user- und projektspezifischen Sicht, wo das automatische Alignment seine Stärken in der Big-Data-Aufarbeitung ausspielen kann und wo die Resultate nicht befriedigen können. Ich zeige, wie wir in unserem Projekt die automatische Segmentation einsetzen, wie wir manuell nacharbeiten, um zu einer Segmentation zu kommen, die unseren Erfordernissen besser entspricht, und wie sich die Resultate von automatischer und manueller Segmentation unterscheiden. Ich zeige auch, wo wir die automatische Segmentation als Basis für eine halbautomatisierte Weiterverarbeitung nehmen können. Das automatische Alignment kann also überzogene Erwartungen leider nicht erfüllen, der Boden der Realität ist aber nicht mehr ganz so hart wie auch schon.

Corpus Phonology - using corpora in phonological and phonetic research

Ulrike Gut

University of Münster
gut@uni-muenster.de

This talk presents Corpus Phonology, a relatively new approach in phonetic and phonological research that makes use of spoken corpora data (Durand, Gut & Kristoffersen 2014). After a definition of what constitutes a phonological corpus and a description of the compilation process, the pros and cons of corpora in phonological and phonetic research will be compared with other types of empirical data, i.e. speech databases and experimentally elicited data.
Subsequently, examples will be given of the use, enhancement and reuse of corpora in phonological and phonetic research. First, the analysis of /h/-deletion and /h/-insertion in Nigerian English will be presented that was based on ICE Nigeria (Wunder et al. 2010), a one-million-word corpus of Nigerian English that contains time-aligned orthographic and some phonological transcriptions and comprises a high number of different speaking styles produced by more than 450 Nigerians.
Second, an example of the enhancement of the ICE Scotland and ICE Nigeria corpora will be given: it will be shown how, in a CLARIN-D project, time-aligned automatic phonological transcriptions generated by the phonemic forced alignment system MAUS (Schiel 2004) were added in order to compare speaker fluency across the two varieties of English (Gut & Fuchs, to appear).
Third, a study on the effect of different learning environments (e.g. stay abroad, training course) on the development of the phonology/phonetics of both L2 German and L2 English will be presented that reuses the LeaP corpus, a 12-hour corpus with 341 recordings by 131 learners of English and of German with a total of 32 different first languages as well as 18 recordings with native speakers of both languages (Gut 2009, 2012). In a conclusion, the advantages and disadvantages of using, enhancing and reusing phonological corpora in a corpus-phonological approach will be discussed.

References

[1] Durand, J., Gut, U. & Kristoffersen, G. (Eds.) (2014). *The Oxford Handbook of Corpus Phonology*. Oxford: Oxford University Press.

[2] Gut, U. (2012). The LeaP corpus. A multilingual corpus of spoken learner German and learner English. In Schmidt Thomas & Wörner Kai (eds.), *Multilingual Corpora and Multilingual Corpus Analysis*, pp. 3-23. Amsterdam: John Benjamins.

[3] Gut, U. (2009). *Non-native speech. A corpus-based analysis of phonological and phonetic properties of L2 English and German*. Frankfurt: Peter Lang.

[4] Gut, U. & Fuchs, R. (to appear). Exploring speaker fluency with phonologically annotated ICE corpora. To appear in *World Englishes*.

[5] Schiel, Florian. (2004). MAUS Goes Iterative. *Proceedings of the IV. International Conference on Language Resources and Evaluation, Lisbon, Portugal*, 1015-1018.

[6] Wunder, E.-M., Voormann, H. & Gut, U. (2010). The ICE Nigeria corpus project: Creating an open, rich and accurate corpus. *ICAME Journal* 34, 78-88.

VORTRÄGE

Ist gleich schnell gleich schnell? Perzeption intendierter und realisierter Sprechrate von deutschen und französischen Sprechern

Zofia Malisz[1,2], Frank Zimmerer[1], Emmanuel Ferragne[3], François Pellegrino[3], Erika Brandt[1], Bistra Andreeva[1], Bernd Möbius[1]

[1]Computerlinguistik und Phonetik, Universität des Saarlandes, Saarbrücken
[2]Deptartment of Speech, Music and Hearing, KTH Stockholm
[3]Laboratoire Dynamique du langage, Université Lyon

Wir stellen eine Replikation und Erweiterung einer Studie von Dellwo und Kollegen [1] vor und untersuchen, inwieweit die intendierte Sprechrate (ISR) und die im Labor messbare Sprechrate (LSR) voneinander abweichen und in welcher Relation sie zu der von den Versuchspersonen perzipierten Sprechrate (PSR) stehen.
Auf Produktionsseite haben Untersuchungen gezeigt, dass Sprecher in der Lage sind, ihre LSR an geforderte ISR anzupassen. So korrelierten beispielsweise schnelle oder sehr schnelle ISR mit entsprechenden LSR (z.B., [2]). Dellwo und Kollegen untersuchten in einem weiteren Schritt, inwieweit Hörer in der Lage sind, die Geschwindigkeitsunterschiede kategorial wahrzunehmen [1]. Ihre Resultate legen nahe, dass französische Hörer in der Lage waren, die ISR verschiedener Muttersprachler einzuordnen. Ein Vorteil für die genauere Einordnung von Sprachstimuli der eigenen Muttersprache konnte ebenfalls gefunden werden.
Unsere Studie zielt darauf ab, diese Untersuchung zu verfeinern. Wir präsentieren Resultate eines 2-Interval-Forced-Choice-Experimentes mit Sprachfragmenten von deutschen und französischen Sprechern. Die Versuchspersonen (Muttersprache: deutsch) hören dabei immer zwei Fragmente (Dauer ca. 1,5 Sek., siehe [3]), die mit unterschiedlichen oder gleichen ISR produziert wurden, deren LSR dabei allerdings ebenfalls variiert und zu Überschneidungen führen kann (beispielsweise eine „normale“ ISR, deren LSR gleich schnell wie eine „schnelle“ ISR ist). Die Aufgabe der Probanden ist, sich zu entscheiden, ob von den zwei gehörten Fragmenten das erste oder das zweite als schneller wahrgenommen wird. Der Aufbau des Experiments folgt einem Versuchsaufbau von [4], deren Resultate die Schwelle der gerade noch wahrnehmbaren Geschwindigkeitsunterschiede bei 5% sehen. Die Fragmente eines jeden Trials sind dabei von einem Sprecher des BonnTempo Corpus gesprochen [5]. Insgesamt werden Satzfragmente von zwei deutschen und zwei französischen Sprechern (jeweils ein Mann und eine Frau pro Sprache) vorgespielt. Den Versuchspersonen werden 220 Stimulipaare präsentiert, die dann innerhalb von 4 Sekunden bewertet werden müssen.
Die Ergebnisse der Studie können einen weiteren Einblick in die Fragestellung werfen, inwieweit ISR mit PSR korreliert, auch wenn die LSR dabei nicht immer unterschiedlich ist. Außerdem können die Resultate mit denen von [1] verglichen werden, um weitere Einblicke in die Universalität der Perzeption von verschiedenen Sprechraten zu gewinnen.

Referenzen

[1] V. Dellwo, E. Ferragne, and F. Pellegrino, "The perception of intended speech ate in English, French, and German by French speakers," in *Proceedings of the Proc. 3rd International Conference of Speech Prosody*, Dresden, Germany, 2006, pp. 101-104.

[2] V. Dellwo and P. Wagner, "Relations between language rhythm and speech rate," in *Proceedings of the 15th International Congress of Phonetic Sciences (ICPhS XV)*, Barcelona, Spain, 2003, pp. 471-474.

[3] J. Koreman, "Perceived speech rate: The effects of articulation rate and speaking style in spontaneous speech," *Journal of the Acoustical Society of America,* vol. 119, pp. 582-596, 2006.
[4] H. Quené, "On the just noticable difference for tempo in speech," *Journal of Phonetics,* pp. 353-362, 2007.
[5] V. Dellwo, I. Steiner, B. Aschenberner, J. Dancovicová, and P. Wagner, "BonnTempo-Corpus and BonnTempo-Tools: A database for the study of speech rhythm and rate," in *Proceedings of the 8th ICSLP*, 2004.

German learners' productions of English sound contrasts: The role of acoustic properties on accent ratings

Nikola Anna Eger, Eva Reinisch

Institute of Phonetics and Speech Processing, LMU Munich
{eger|evarei}@phonetik.uni-muenchen.de

Learning a second language (L2) requires, among other skills, the mastery of producing non-native sound contrasts. However, even when learners do produce contrasts, the acoustic cues may differ from native speakers' patterns and productions may still sound accented to native speakers. The aim of this study was to investigate the relation between acoustic properties that learners produce to indicate difficult L2 contrasts and native speakers' perception of a foreign accent. Specifically, we focused on German learners of English producing word-final voicing contrasts, and the difficult vowel contrast / /-/æ/. Voicing contrasts do occur in German ɛ but not in word-final position. In wordfinal position voiced obstruents are an unfamiliar category. For perception it has been shown that knowledge about contrasts that exist in some positions can be transferred to unfamiliar positions (Broersma, 2005). The English vowel contrast /ɛ/-/æ/ falls into one single category /ɛ/ in German and is thus an entirely unfamiliar contrast to Germans. It may hence be even more difficult to acquire than the voicing contrast. The present study will address potential effects of contrast familiarity in production and the relation to perceived accent. We predict that speakers with acoustically larger contrasts with more native-like cues will receive better accent ratings than speakers who produce less contrast. This effect should be especially salient when looking at the unfamiliar category /æ/ and the voiced obstruents.
24 German learners of English produced 31 minimal pairs containing these contrasts. For the final voicing contrast we measured among other cues duration of the preceding vowel, consonant duration and voicing. For the vowels we took spectral measures (F1,F2) and duration. For each type of contrast speakers were assigned to one of three proficiency groups according to how well they had produced the contrast, taking into account the cues they had used: The larger the difference between the categories and the less overlap between them, the more "proficient" was the speaker. In a subsequent perception experiment, native speakers of American English rated the goodness of the words from the minimal pairs.
For pairs differing in /ɛ/-/æ/, words with the category /æ/ were rated as clearly more accented than words with /ɛ/. Despite improvement over speaker proficiency group, even speakers with the largest acoustic contrasts differed from ratings for /ɛ/. The voicing pairs show a different pattern: Speakers with good contrasts (i.e., those in the "high proficient group") received equally good or even better ratings for words containing the unfamiliar category. Differences in producing different types of non-native contrasts are discussed in relation to consequences for perceivable accent.

References

[1] Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. The Journal of the Acoustical Society of America, 117(6), 3890-3901.

L3 phonology - cross-language, psychoacoustic, and extralinguistic factors

Christina Domene Moreno, Baris Kabak

University of Würzburg
christina.domene-moreno@uni-wuerzburg.de, baris.kabak@uni-wuerzburg.de

Various factors have been shown to be crucial in determining the source and directionality of crosslinguistic influence (CLI) in Third Language Acquisition (L3A). While lexical (Angelovska & Hahn 2012) and morphosyntactic (Tsang 2009) development in L3A are relatively well documented, there are still comparatively few studies on L3 phonology (e.g., Kopeckova 2015). Furthermore, while some work is appearing on the role of extralinguistic factors on L3 acquisition (e.g., metalinguistic awareness, Wrembel 2015), L3A studies do not yet systematically tease apart CLI from the contribution of multilingualism. This gap begs for new empirical studies in L3A, especially considering the previously shown connection between cognitive (Bialystok et al. 2012) and linguistic (Jessner 2006) advantages of multilingualism.

Our study investigates the nature and extent of transfer, and attempts to establish whether there is a bilingual advantage in the acquisition of an L3 phonology, by separating transfer from non-transfer conditions. In particular, we test young bilingual (Turkish-German) learners of English on their acquisition of segmental (e.g., interdental fricatives) and suprasegmental (e.g., word initial consonant clusters) features in the target language. These features are further grouped into three sets based on the predicted crosslanguage transfer: (A) those promoting positive transfer from the L1 of the bilingual learners (Turkish), (B) those promoting positive transfer from the respective L1/L2 (German), and (C) those that are not expected to provoke transfer from either language.

In our experiment we tested the perception and production of (A), (B), and (C) by Turkish-German learners of English (n=12, mean age=11.5), who all attended grade 5 or 6 at the same German public high school at the time of testing, via an AX discrimination task using nonsense words, and a delayed repetition task using novel compound words. We further investigated whether the perception and production results are modulated by cognitive and sociolinguistic (attitudinal/motivational) variables through a phonological working memory task and questionnaires.

Based on the Cumulative Enhancement Model (Berkes & Flynn 2012), we predicted the bilinguals to perform equally well in (A) and (B). If bilinguals enjoy the cognitive and linguistic benefits of bilingualism, there should be a positive correlation between the phonological memory scores and the ability to perceive and produce all English contrasts in question.

Our results so far do not point to any distinct ease with in any one of the two categories of contrasts from their background languages (i.e., A and B). Rather, phonological transfer appears to be cumulative in our bilinguals. Furthermore, the perception and production of the features tested is non-isomorphic; what appears to be best perceived is not necessarily the best produced albeit the substitution of interdental fricatives being perceptually conditioned with no reflection of the typical substitution patterns of German or Turkish. Finally, phonological memory is not a significant factor for success in either production or perception of the tested features, although a comparison with an age- and SES- matched monolingual German group is necessary in order to clearly tease apart the contribution of extralinguistic factors. While we argue that the bilingual population should be studied in its own right, we will present results from a study (currently underway), where we also tested age- and SES-matched monolingual German learners of English on the same contrasts in order to establish acquisition hierarchies for L2 and L3 phonological learning.

This comparison will especially be instructive to investigate whether the bilingual values for English contrasts converge with monolingual values for the same contrasts in production.

References

[1] Angelovska, T., & Hahn, A. (2012). Written L3 (English) Transfer Phenomena of L2 (German) Lexical and Syntactic Properties. In D. Gabrys-Barker (Ed.), *Cross-Linguistic Influences in Multilingual Language Acquisition* (pp. 23-40). Berlin: Springer.

[2] Berkes, É., & Flynn, S. (2012). Further evidence in support of the Cumulative-Enhancement Model: CP structure development. In J. Cabrelli Amaro, S. Flynn, & J. Rothman (Eds.), *Third Language Acquisition in Adulthood* (pp. 143-164). Amsterdam: John Benjamins.

[3] Bialystok, E. et al. (2012). Bilingualism: consequences for mind and brain. *Trends in Cognitive Sciences*, 16 (4), 240-250.

[4] Jessner, U. (2006) *Linguistic awareness in multilinguals: English as a Third Language.* Edinburgh: Edinburgh University Press.

[5] Kopeckova, R. (2015). Differences in the perception of English vowel sounds by child L2 and L3 learners. In Gut, U., Fuchs, R. & Wunder E.-M. (Eds.), *Universal or Diverse Paths to English Phonology?* (pp. 71-89). Berlin: Mouton De Gruyter.

[6] Llama, R. et al. (2010). The roles of typology and L2 status in the acquisition of L3 phonology: The influence of previously learnt languages on L3 speech production. *International Journal of Multilingualism*, 7 (1), 39-57.

[7] Tsang, W. L. (2009). The L3 acquisition of Cantonese reflexives. In Y-k. I. Leung (Ed.), *Third Language Acquisition and Universal Grammar* (pp.192-219). Bristol: Multilingual Matters.

[8] Wrembel, M. (2015). *In search of a new perspective: Cross-linguistic influence in the acquisition of an L3 phonology*. Poznan: Adam Mickiewicz University Press

Ein computer-gestützter Vokaltrainer für die deutsche Sprache

Frank Zimmerer, Jeanin Jügler

Institut für Computerlinguistik und Phonetik, Universität des Saarlandes
zimmerer@coli.uni-saarland.de, juegler@coli.uni-saarland.de

Der Erwerb von Vokalen in einer Fremdsprache (L2), sowohl im Bezug auf Produktion als auch auf Perzeption, ist für Sprecher verschiedener Muttersprachen kein einfaches Unterfangen, u.a. [1, 2, 3, 4, 5, 6, 7]. Das Vokalsystem des Deutschen erscheint für viele Lerner komplex. Grund hierfür ist einerseits die relativ große Anzahl von 16 Monophthongen. Andererseits, und damit verbunden, unterscheidet das Deutsche zwischen gespannten und ungespannten Vokalen und zusätzlich zwischen langen und kurzen. Typologisch markiert sind ebenfalls vordere, gerundete Vokale (/y: ʏ, o, oe/).
Ergebnisse verschiedener Analysen des IFCASL Korpus [8] (gelesene Sprache) haben gezeigt, dass französische Lerner des Deutschen Probleme in der korrekten Produktion von Vokalen bezüglich Lange und Qualität haben [9, 7]. Beispielsweise produzieren einige Lerner in dem Satz „Im Frühling fliegen Pollen durch die Luft.“, das ungespannte /ɔ/ als ein gespanntes [o:]. Außerdem, teilweise auf Orthographie---Aussprachеinterferenzen beruhend, wird <u> auch als [y:, ʏ] produziert. Lerner mit anderen muttersprachlichen Hintergründen zeigen ebenfalls Probleme im L2-Erwerb von Vokalen, jeweils im Zusammenspiel des phonologischen Vokalsystem der Muttersprache mit dem
deutschen Vokalsystem.
Der Vokaltrainer soll Lernern des Deutschen mit unterschiedlichen Muttersprachen helfen, die Produktion der Vokale hinsichtlich Qualität und Quantität zu üben und zu erlernen. Zuerst fordert das Programm die Lerner zur Produktion der Kardinalvokale /a:, i:, u:/ auf, um die Formantwerte der jeweiligen Sprecher zu kalibrieren. Anschließend können sie einen oder zwei Vokale auswählen, um diese zu üben. Im Übungsmodus bietet das Programm die Möglichkeit, Beispielwörter (und Pseudo-Wörter) von deutschen Muttersprachlern zu hören und die
Aussprache zu trainieren. Die Zielvokale werden durch Seerosen dargestellt. Deren Positionierung ist an die F1/F2---Verteilung im Vokalviereck angelehnt. Ziel für die Lerner ist es, einen Frosch durch die korrekte Produktion der Wörter mit dem Vokal auf die Seerose springen zu lassen (das heißt, dass die Formantwerte der Produktion mit denen der Referenzsprecher hinreichend übereinstimmen). Wenn die Produktion falsch ist, fällt der Frosch ins Wasser. Die korrekte Produktion der Vokallänge wird durch Balken oberhalb der Seerose angezeigt. Das Programm protokolliert den Fortschritt der Benutzer und zeigt durch die Größe der Blüte der Seerose, inwieweit die Produktion des jeweiligen Vokals als muttersprachlich erachtet werden kann.
So soll spielerisch das Üben der Vokale überwacht werden, die Lerner sollen leicht verständliches Feedback erhalten, und die Protokollfunktion soll einen Lernfortschritt erkennbar machen. Implizit soll so auch die korrekte Zungenposition für Vokale im Deutschen erlernt werden.

Referenzen

[1] I. Darcy and F. Kruger, “Vowel perception and production in Turkish children acquiring L2 German,” Journal of Phonetics, vol. 40, pp. 568–581, 2012.
[2] J. E. Flege and J. M. Hillenbrand, “Limits on phonetic accuracy in foreign language speech production,” Journal of the Acoustical Society of America, vol. 76, no. 3, pp. 708–721, 1984.

[3] J. E. Flege and I. R. A. MacKay, "Perceiving vowels in a second language," Studies in Second Language Acquisition, vol. 26, pp. 1–34, 2004.
[4] J. Kingston, "Learning foreign vowels," Language and Speech, vol. 46, no. 2-3, pp. 295–349, 2003.
[5] J.-W. Van Leussen and P. Escudero, "Learning to perceive and recognize a second language: the l2lp model revised," Frontiers in Psychology, vol. 6, no. 1000, 2015.
[6] W. Strange, O.-S. Bohn, S. A. Trent, and K. Nishi, "Acoustic and perceptual similarity of North German and American English vowels," Journal of the Acoustical Society of America, vol. 115, pp. 1791–1807, 2004.
[7] F. Zimmerer and J. Trouvain, "Perception of French speakers' German vowels," in Interspeech 2015, 2015, pp. 1720–1724.
[8] J. Trouvain, A. Bonneau, V. Colotte, C. Fauth, D. Fohr, D. Jouvet, J. Jugler, Y. Laprie, O. Mella, B. Möbius, and F. Zimmerer, "The ifcasl corpus of French and German non-native and native read speech," in Proc.10th Language Resources and Evaluation Conference (LREC), Portoroz (Slovenia), 2016, pp. 1333–1338.
[9] D. Jouvet, A. Bonneau, J. Trouvain, F. Zimmerer, Y. Laprie, and B. Möbius, "Analysis of confusion matrices in a manually annotated French-German learner corpus," in Workshop on Speech and Language Technology in Education (SlaTE 2015), Leipzig (Germany), 2015, pp. 107–112.

PATSY-I: A Corpus on Non-Native English Air Traffic Communication

Caroline Kaufhold[1], Christine Martindale[1], Axel Horndasch[1]
Klaus Reinhard[2], Elmar Nöth[1]

[1]Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nuremberg, Germany
[2]e.sigma Technology GmbH, Ilmenau, Germany
{caroline.kaufhold|christine.f.martindale|axel.horndasch|elmar.noeth}@fau.de,
kreinhard@esigma-technology.com

In many global tasks English is used as an international language. As a consequence, non-native speakers of the English language often communicate with other non-native speakers. An example is the Air Traffic Control (ATC) service which directs aircrafts on the ground and through controlled airspace. It is of course essential that there is a perfect understanding between the pilot and the ground-based controller to prevent collisions and to organize air traffic efficiently. Aviation English already accommodates non-native speakers of English by providing guidelines for wording and phraseology. To avoid confusion, for example, letters and numbers are spelled according to the international spelling alphabet provided by the International Civil Aviation Organization (ICAO). However, the ability to speak and understand English still has a high impact on communication success. In this paper, we present a corpus that was recorded in the context of the ATC phraseology training system PATSY, the prototype of which was presented at the Show&Tell session at Interspeech 2015 [1]. The corpus consists of basic ATC utterances by speakers of 16 different mother tongues. Furthermore, “Please Call Stella” [2] and part of “The Rainbow Passage” [3] were recorded twice for every speaker with different biases. We plan on using this data to study the entrainment effect, which was observed for conversations by Levitan and Hirschberg [4]. Preliminary results on basic ATC utterances show a moderate correlation between the speakers’ self-assessment and the GoP (Goodness of Pronunciation) score.

☞ **See full paper in Proceedings**

Recording a High-Quality German Speech Database for the Study of Speaker Personality and Likability

Laura Fernández Gallardo

Quality and Usability Lab, Telekom Innovation Laboratories, TU Berlin, Germany
laura.fernandezgallardo@tu-berlin.de

The ongoing process for recording a personality and likability database in German is motivated and described. Overall, highquality and consistency among recordings is pursued, in order to avoid possible biases when rating speaker characteristics and low performance when automatically detecting them. Prescribed and spontaneous human-human dialogs are recorded using three different microphones in an acoustically-isolated room with 48 kHz sampling frequency. Natural and neutral speech is recorded, controlling the absence of background noises. So far, 101 German speakers without accent have participated and two listening tests have been conducted with part of these data. Our goal is to extend the number of recorded participants to at least 200, conduct more listening tests, and share the speech files, metadata, and associated labels, features, and analyses with the scientific community.

☞ **See full paper in Proceedings**

Phonetic learner corpora for research on German as a foreign language

Jürgen Trouvain, Frank Zimmerer

Saarland University, Saarbrücken
{trouvain|zimmerer}@coli.uni-saarland.de

The use of phonetic learner corpora is rather new in the fields of teaching a foreign language (L2), the research of L2 acquisition, and studies of non-native speech in general. This holds in particular for German as a foreign language (Deutsch als Fremdsprache or DaF) where phonetic topics are either handled with the development of teaching material or, more rarely, on experimental studies. The teaching material is usually either based on long-term subjective experience of teaching phonetic topics in the classroom, or on contrastive comparisons of first and foreign languages. The few experimental phonetic studies with German as L2 concentrate on selected phenomena with a limited number of subjects recorded under highly controlled conditions.
In contrast to that, corpora enable analyses of presumed interferences with a large number of learner data. Phonetic learner corpora focus on phonetic and phonological questions considered in the design and the (often automatically performed) annotation. The audio and annotation files make it possible to perform detailed acoustic and auditory analysis of numerous important and relevant interferences.
With our French-German phonetic learner corpus [1] we illustrate selected segmental and prosodic phenomena in the production of read speech by French learners of German. These include e.g. for the consonants how /h/ is realised compared to German native speakers which often includes a glottal stop as a surrogate (in contrast to an often assumed deletion of /h/) [4]. A notoriously difficult field, not only for French speakers, is the realisation of the appropriate quantity and quality of German vowels. The analysis of annotations and subsequent perception tests using minimal pair words from the corpus show a particular problem with rounded vowels (not predictable from contrastive analysis alone) and huge individual differences across proficiency levels [5]. An example for the analysis of a prosodic problem is a study of the pauses and disfluent phases of a read text which delivers important background on the speech fluency of learners [2].
Further examples of how to use phonetic learner corpora are the development of tools for computer-assisted pronunciation training [3] as well as using the corpus data for illustrations and the generation of exercises with real audio data. The first application addresses DaF learners whereas the latter is of interest for (prospective) DaF teachers.

References

[1] Trouvain et al. 2016: The IFCASL corpus of French and German non-native and native read speech. Proc. LREC.
[2] Trouvain, Fauth & Möbius 2016. Breath and non-breath pauses in fluent and disfluent phases of German and French L1 and L2 Read Speech. Proc. Speech Prosody.
[3] Vakil, A. 2015. A CAPT tool for training and research on lexical stress errors in German. Proc. SLaTE.
[4] Zimmerer & Trouvain 2015a. Productions of /h/ in German: French vs. German speakers. Proc. Interspeech.
[5] Zimmerer & Trouvain 2015b. Perception of French speaker's German vowels. Proc. Interspeech.

Big Data for analyses of small-scale regional variation: A case study on sound change in Swiss German

Adrian Leemann[1], *Marie-José Kolly*[2]

[1]Phonetics Lab., Department of Theoretical and Applied Linguistics, University of Cambridge
[2]Phonetics Lab., Department of Comparative Linguistics, University of Zurich
al764@cam.ac.uk, marie-jose.kolly@uzh.ch

In this case study we examine sound change of *Altoberdeutsch* <iu> in Swiss German dialects. We used contemporary dialect data from nearly 60,000 speakers – collected with the smartphone app *Dialäkt Äpp* – and compared it to historical *Atlas* data from the 1950s. Results revealed hierarchical and contra-hierarchical diffusion patterns for some dialectal variants, while other variants remained virtually unchanged over the course of seven decades. We further report change in apparent time, with older speakers using traditional variants more frequently than younger speakers. Using this case study as a model, future work using the *Dialäkt Äpp* corpus will reveal patterns of feature diffusion and dialect leveling on a larger scale.

☞ **See full paper in Proceedings**

Lexical and prosodic operators to contrastive alternatives

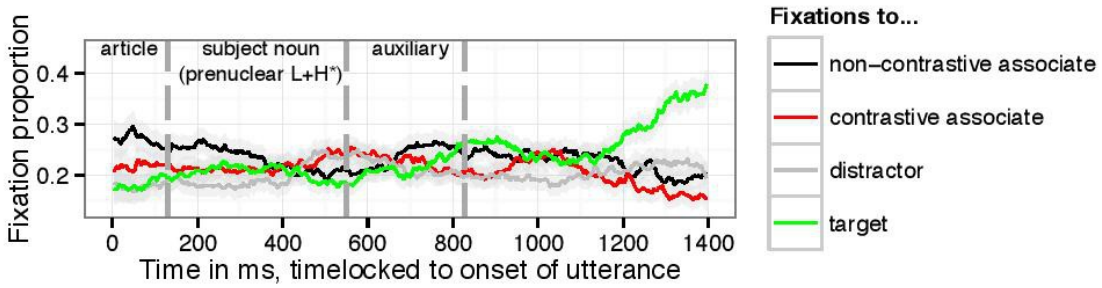
Bettina Braun, Yuki Asano & Nicole Dehé

Universität Konstanz
{barbara.werner|yuki.asano|nicole.dehe}@uni-konstanz.de

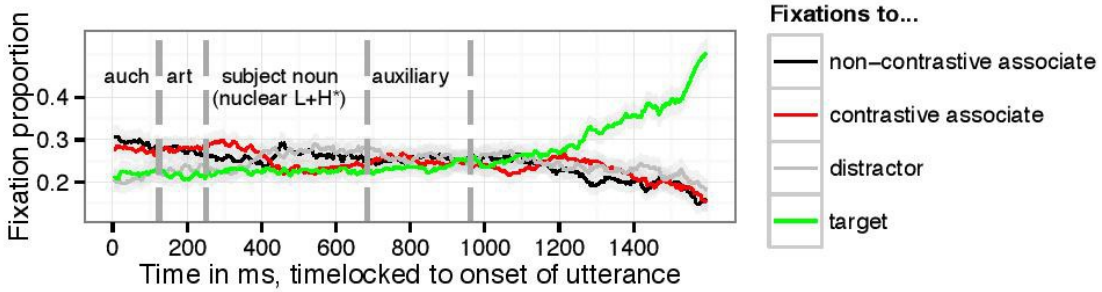
Recent research has shown that narrow focus constituents with contrastive pitch accents (e.g., nuclear L+H* L- in German, cf. Kügler and Gollrad (2015), Baumann, Grice, and Steindamm (2006)) lead the activation of alternatives (Braun & Tagliapietra, 2010; Husband & Ferreira, 2012). Furthermore, contrastive alternatives are recalled better when the target was contrastively accented (Fraundorf, Watson, & Benjamin, 2010; Gotzner, 2014; Gotzner, Spalek, & Wartenburger, 2013). Contrastive alternatives are not only triggered by contrastive pitch accents but also by certain lexical particles, like additive particles (*auch* 'also'): Gotzner and colleagues reported that recall of contrastive alternatives is facilitated with additive particles (*auch* 'also), but that exclusive particles (*nur* 'only') hinders recall (Gotzner & Spalek, 2014; Spalek, Gotzner, & Wartenburger, 2014). It is largely unclear, however, how contrastive referents are activated as the utterance unfolds over time. Braun (2015) used the visual world eye tracking paradigm to test whether listeners fixate more on referents that are contrastively related to the subject of an auditory sentence (henceforth contrastive associate) when the subject was realized as a narrow focus (either with a contrastive accent: nuclear L+H* L- or a non-contrastive accent: nuclear H+L* L-) compared to a broad focus realization (prenuclear L+H* accent on the subject noun). Only a contrastive accent on the subject (in e.g., *Der Schwimmer wollte Flossen anziehen*) led to more fixations to the contrastive associate (e.g., *diver*).
Here, we investigate the integration of focus sensitive particles (*also, only*) with the prosodic realization of the subject noun (contrastive vs. non-contrastive accent). These focus sensitive particles associate with the accented subject noun. The materials and procedure were as described in Braun (2015): Participants saw a visual display with four printed words, one being the contrastive associate to the subject in the auditory sentence, one being non-contrastively related to it, one the object of the sentence (which had to be clicked) and an unrelated distractor. In Experiment 1a and 1b, the experimental sentences started with the additive particle *also* (*Auch der Schwimmer wollte Flossen anziehen*), once with a contrastive nuclear L+H* L- on the subject, once with a noncontrastive nuclear H+L* L- on the subject. The control condition was a broad focus realization without particle (*Der Schwimmer wollte Flossen anziehen*). In Experiment 2, the sentence started with the exclusive particle *only* (*Nur der Schwimmer wollte Flossen anziehen*). Here, we only have data for the contrastive accent and the control condition. Each experiment had 24 experimental trials (12 for each condition) and 24 filler trials (in which the display showed contrastive associates to the object noun). For Experiment 1a and 1b, we tested 40 German native speakers, students at the University of Konstanz, for Experiment 2, 27 participants. We used the SMI Eyelink 1000 system. Fixations were extracted in 4ms steps and coded as pertaining to one of the four words on screen.
The evolution of fixations to the words is shown in Figure 1 for all experiments (fixations to the contrastive associate are shown in red). For statistical analysis, empirical logits of fixations to the contrastive associate were calculated and compared for each experimental condition to the respective control condition (Barr, Gann, & Pierce, 2011). Empirical logits were analyzed using linear-mixed effects regression models with condition as fixed factor and random intercepts and slopes for participants and items (Baayen, Davidson, & Bates, 2008; Baayen, 2008; Barr, Levy, Scheepers, & Tily, 2013). Listeners fixated the contrastive associate more when the sentence started

with *auch* and the subject noun was produced with a contrastive nuclear L+H* L- accent compared to the control condition ($p < 0.05$). No other comparisons were significant. There is an intriguing interplay between focus-sensitive particles and contrastive accents. In sentences with the additive particle, the accent type of the subject noun matters: only a contrastive accent on the subject noun led to the activation of alternatives, similar to sentences without particle (Braun, 2015), so contrastive accents outweigh lexical information. On the other hand, a contrastive accent following an exclusive particle did not lead to the activation of alternatives: the lexical semantics seems to outweigh prosodic contrast marking.

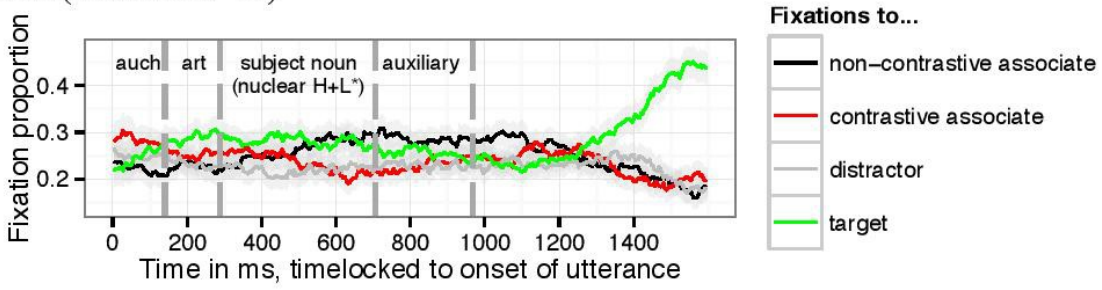
Fig 1. Evolution of fixations in the different experiments, time-locked to the onset of the sentence
- control condition (prenuclear L+H* L-):



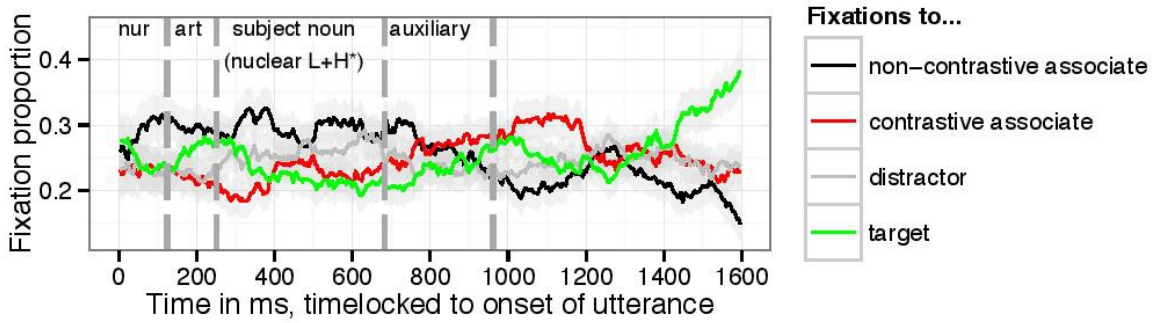
- auch (nuclear L+H* L-):



- auch (nuclear H+L* L-):



- nur (nuclear L+H* L-):



References

[1] Baayen, H. R., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390-412.

[2] Baayen, R. H. (2008). *Analyzing Linguistic Data. A Practical Introduction to Statistics*. Cambridge: Cambridge University Press.

[3] Barr, D. J., Gann, T. M., & Pierce, R. S. (2011). Anticipatory baseline effects and information integration in visual world studies. *Acta Psychologica, 37*, 201-207.

[4] Barr, D. J., Levy, R., Scheepers, C., & Tily, H. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 36*, 255-278.

[5] Baumann, S., Grice, M., & Steindamm, S. (2006). *Prosodic marking of focus domains – categorical or gradient?* Paper presented at the 3rd International Conference on Speech Prosody, Dresden, Germany.

[6] Braun, B. (2015). *What causes the activation of contrastive alternatives: The size of focus domain or pitch accent type*. Paper presented at the 18th International Congress of the Phonetic Sciences.

[7] Braun, B., & Tagliapietra, L. (2010). The role of contrastive intonation contours in the retrieval of contextual alternatives. *Language and Cognitive Processes, 25*, 1024-1043.

[8] Fraundorf, S. H., Watson, D. G., & Benjamin, A. S. (2010). Recognition memory reveals just how CONTRASTIVE contrastive accenting really is. *Journal of Memory and Language, 63*, 367-386.

[9] Gotzner, N. (2014). *Establishing alternative sets.* Humboldt-Universität zu Berlin.

[10] Gotzner, N., & Spalek, K. (2014). *Exhaustive inferences and additive presuppositions: The interplay of focus operators and contrastive intonation.* Paper presented at the Proceedings of the ESSLLI workshop on formal and experimental pragmatics.

[11] Gotzner, N., Spalek, K., & Wartenburger, I. (2013). *How pitch accents and focus particles affect the recognition of contextual alternatives.* Paper presented at the Proceedings of the 35th Annual Conference of the Cognitive Science Society, Austin, TX.

[12] Husband, M. E., & Ferreira, F. (2012). *Generating contrastive alternatives: Activation and suppression mechanisms.* Paper presented at the 25th Annual CUNY Conference on Human Sentence Processing, New York.

[13] Kügler, F., & Gollrad, A. (2015). Production and perception of contrast: The case of the rise-fall contour in German. *Frontiers in Psychology, 6*.

[14] Spalek, K., Gotzner, N., & Wartenburger, I. (2014). Not only the apples. Focus-sensitive particles improve memory for information-structural alternatives. *Journal of Memory and Language, 70*, 68-84.

The impact of animacy and rhythm on the word order of conjuncts in German

Isabelle Franz [1], *Gerrit Kentner* [1], *Frank Domahs*[2]

[1] Goethe - Universität Frankfurt
[2] Philipps - Universität Marburg
i.franz@emuni-frankfurt.de, kentner@lingua.uni-frankfurt.de, domahs@uni-marburg.de

In this study we investigated the impact of two constraints on the linear order of constituents in children's speech production. Two types of constraints have been found to influence serialization in English speaking participants: a rhythmic (*LAPSE) and a semantic (ANIM) one. We tested 18 German children aged three to six years. Participants were instructed to produce coordinated bare noun phrases in response to picture stimuli (e.g., *Delphin und Planet)*. Disyllabic target words were controlled with respect to word stress and animacy.
Overall, children preferably produced animate items before inanimate ones, confirming findings of Prat-Sala, Shillcock and Sorace (2000) [1]. Furthermore, the order of the conjuncts was affected by the rhythmic constraint, such that disrhythmic constructions (resulting in a sequence of unstressed syllables, a so called stress lapse) were avoided. The latter results were significant when the factor animacy didn't vary and can be taken as evidence for the *prosodic licensing hypothesis* (Demuth 2007) [2]. In sum, our findings suggest a stronger influence of animacy compared to rhythmic well-formedness on conjunct ordering for German speaking children, as it was shown for English speaking adults by McDonald and colleagues (1993) [3].

☞ **See full paper in Proceedings**

Anticipatory V-to-V Coarticulation in German Preschoolers

Elina Rubertus [1], *Dzhuma Abakarova* [1], *Jan Ries* [1], *Aude Noiray* [1]

[1] Laboratory for Oral Language Acquisition, University of Potsdam
rubertus@uni-potsdam.de

This study investigates lingual V-to-V anticipatory coarticula-tion in German preschoolers and adults using ultrasound measures. In light of conflicting results in the literature, the aim was to study effects in larger cohorts and with a widespread set of vowels. Results provide evidence for V-to-V coarticulation in children as well as adults, independent of the intervocalic consonant. Interestingly, coarticulation degree decreases with age.

☞ **See full paper in Proceedings**

On the relationship between pointing gestures and speech production in German counting out rhymes: Evidence from motion capture data and speech acoustics

Susanne Fuchs[1], *Uwe D. Reichel*[2]

[1] Zentrum für Allgemeine Sprachwissenschaft, Berlin, Germany
[2] Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, Hungary
fuchs@zas.gwz-berlin.de, uwe.reichel@nytud.mta.hu

We investigated the interplay between pointing gestures and speech by means of motion capture and acoustics. Counting out rhymes served as a testbed, since they involve clear index finger turning points. The distance between the participant and an interlocutor (a teddy) was varied between close and far. Additionally, speaking with a normal speech rate in comparison to a fast rate was examined. Results of 1352 pointing gestures provide evidence that: a) the number of syllables realized per stroke are in general relatively stable across condition, but differences occur among subjects, b) turning points occur frequently in vowels, but also in consonants when syllables have a complex phonological structure, c) fast speech rate not only affects speech, but also leads to a shortening in pointing gesture duration, d) rhythmicity of the strokes is reduced with high speech rate and e) the impact of stroke rate on the acoustic energy contour is larger in normal than in fast speech. Distance showed no strong effects. We believe that counting out rhymes show great potential for further research in which further insight could be gained into the rhythmic and prosodic characteristics of a language as well as the coordination between pointing gestures and speech.

☞ **See full paper in Proceedings**

(Mor-)phonotactic consonant clusters in Standard Austrian German and Standard German German

Hannah Leykum[1], *Sylvia Moosmüller*[1]

[1] Acoustics Research Institute, Austrian Academy of Sciences, Austria
hannah.leykum@oeaw.ac.at, sylvia.moosmueller@oeaw.ac.at

Consonant clusters occur within morphemes (phonotactic clusters) as well as across morpheme boundaries (morphono-tactic clusters). Since morphonotactic clusters contain mor-phological information, differences between the two types of clusters are expected in speech production. Previous studies on language acquisition, speech processing, and computer simu-lations proved a different treatment of these types of clusters.
Our previous analyses of Standard Austrian German (SAG) speakers showed no significant difference in the production of the two types of clusters. We interpreted this as a result of specific timing relations in SAG, which might impede a different treatment of the two types of clusters. In order to prove this hypothesis, we compared word-final phonotactic and morphonotactic consonant clusters in homo-phonous word-pairs produced by speakers of SAG, speakers of Standard German German living in Germany (SGG), and speakers of Standard German German living in Austria (SGGA).
The analyses revealed that, as expected, the speakers of SAG did not differentiate between the two types of clusters. Whereas, the speakers of SGG and SGGA as well, did no differentiation between phonotactic and morphonotactic clusters in speech production. Therefore, the hypothesis on an influence of the prosody could not be confirmed.

☞ **See full paper in Proceedings**

The Acoustics of Fricative Contrasts in Two German Dialects

Stefanie Jannedy [1], *Melanie Weirich*[2]

[1] ZAS Berlin, Germany
[2] Friedrich-Schiller Universität Jena, Germany
jannedy@gmail.com, melanie.weirich@uni-jena.de

In this study, we are investigating the acoustic characteristics of the five voiceless German fricatives [f s ç ʃ χ] elicited in non-words from 3 speakers each of two German dialects. The northern German dialect differentiates these five fricatives whereas in the middle German region and in Berlin, /ç/ and /ʃ/ have already merged or are in the process of merging [1;2]. Our previous work (submitted) has indicated that differentiating [ç] and [ʃ] acoustically in the speech of northern speakers from that of Berlin speakers works best when using DCT (discrete cosine transformation) rather than the four spectral moments. Results from our study corroborate this finding for both the northern and the middle German dialect.

☞ **See full paper in Proceedings**

Monolingual and trilingual production of Northern Standard German vowels

Heike Schoormann[1], *Wilbert Heeringa*[2], *Jörg Peters*[1]

[1] Institute for German Studies, University of Oldenburg, Oldenburg, Germany
[2] Fryske Akademy, Ljouwert, the Netherlands
heike.schoormann@uol.de, wjheeringa@gmail.com, joerg.peters@uol.de

Studies on vowel productions of speakers from bilingual communities report L1-L2 interactions but also monolinguallike realizations ([1], [2], [3]). Where the languages differed in communicative range and size of the speech community, monolingual-like productions of early bilinguals were found in the languages with the wider communicative range and larger speech community. We compare the acoustic realizations of Northern Standard German (NSG) vowels in monolingual speakers from Hanover, representing the larger speech community of Northern Germany, and in trilingual speakers from the Saterland, speaking the local variant of High German, Low German, and Saterland Frisian. To examine whether the NSG vowels of the Saterland speakers approached the vowels of the monolingual speakers in terms of spectral and durational features, we elicited all stressed NSG monophthongs in /hVt/ context. Our data show an orientation towards the larger speech community of Northern Germany in the productions of the trilinguals. Vowel productions which neither differed across the trilinguals' three languages nor from the monolinguals suggest contact-induced phonetic convergence towards NSG. The observed bidirectional interaction of the trilinguals' three vowel systems further supports the claim that all vowel categories are organized in a common phonological vowel space.

☞ **See full paper in Proceedings**

Sprachdatenerhebung und Kontextvariation: Wie die Intonation in Fragen vom kommunikativen Kontext gesteuert wird

Benno Peters [1], *Matthias Hoffmann* [2], *Laura-Marie Andresen* [3]

ISFAS-Abteilung für Allgemeine Sprachwissenschaft, CAU Kiel [1,2,3]
peters@ipds.uni-kiel.de, stu107929@mail.uni-kiel.de, laura.andresen32@web.de

Phonetische Strukturen von Äußerungen unterscheiden sich in Abhängigkeit von Gesprächssituationen und interpersonellen Beziehungen. So können sich Gesprächspartner ihrem Gegenüber unterordnen oder sich in die dominante Position begeben. Eine solche Positionierung zeigt sich unter anderem in der Intonation von Fragen. Die systematische empirische Untersuchung solcher Phänomene braucht eine geeignete Datengrundlage. Wir benötigen Sprachdaten aus Situationen, in denen sich Gesprächspartner interaktional klar positionieren. Die Methode gut analysierbare Zielsätze in interaktional steuernde Kontexte einzubetten ist geeignet, um systematisch Daten in gelesener Sprache zu erheben, die ebenfalls Aufschlüsse über Alltagskommunikation geben können. Folgendes Beispiel zeigt wie der Zielsatz *Was sollen wir machen?* (a) in einem sich unterordnenden Kontext eingebettet ist: *Ich weiß nicht weiter. Was sollen wir machen?* und (b) auf einen initialen Kontext folgt, der Dominanz signalisiert: *Na toll. Du hast das kaputt gekriegt. Was sollen wir machen?* Eine solche Einbettung resultiert in sich klar unterscheidenden Verteilungen phrasenfinaler Intonationsmuster in den verschiedenen Kontexten. Die Satzintonation fällt typischerweise in dominanten Fragen und steigt in sich unterordnenden Fragen. Dieser Beitrag stellt die Funktionsweise und das Potential der Methode der Kontextvariation am Beispiel der Untersuchung der phrasenfinalen Intonation in W-Fragen im Deutschen dar.

☞ **See full paper in Proceedings**

Prosodic Marking of Information Status in Task-Oriented Dialogues

Christine T. Röhr, Tabea Thies, Stefan Baumann, Martine Grice

I*f*l Phonetik, University of Cologne, Germany
{christine.roehr|tabea.thies|stefan.baumann|martine.grice}@uni-koeln.de

In the present paper we investigate the effect of information status on accent placement and accent types used in semispontaneous speech. As an elicitation method we use a 'spotthe-difference' task which provides natural (dialogue) but still controlled (task-oriented) speech data. The task has been shown to be an ideal testbed for the relation between prosody and discourse meaning. However, it has not been used in a fine-grained study of information status yet. This is done in the present study by applying the *RefLex* annotation scheme, which differentiates between a referential and a lexical level of givenness. The semi-spontaneous speech data indicate a systematic but probabilistic relation between prosodic prominence and an item's level of givenness. That is, the correlation between increasing newness and increasing prominence is predominantly reflected in a more frequent use of nuclear pitch accents as well as a less frequent use of deaccentuation. Both the referential and lexical levels of givenness turn out to have an incremental effect on the degree of an item's prosodic prominence. Consequently, the *combined* degree of givenness of a referent (reflected by a combination of RefLex labels) indicates an overall prominence value of the item's prosodic realization.

☞ **See full paper in Proceedings**

Teasing apart lexical stress and sentence accent in Hungarian and German

Ádám Szalontai[1], Petra Wagner[2], Katalin Mády[1], Andreas Windmann[2]

[1]Research Institute for Linguistics, Hungarian Academy of Sciences
[2]University of Bielefeld
{szalontai.adam|mady.katalin}@mta.nytud.hu, {petra.wagner|awindmann2}@uni-bielefeld.de

This study compares the strategies to mark lexical stress and sentence-level accent in Hungarian and in German by employing two production experiments of comparable designs. The experimental conditions elicited target segments in +/– stressed and +/– accented conditions. The results indicated that while German, a language with variable lexical stress placement, clearly marks both stress and accent with a number of phonetic parameters, Hungarian, a language with fixed word-level stress placement marks accents, but not stress.

☞ **See full paper in Proceedings**

DIMA – Erweiterte Richtlinien zur Annotation deutscher Intonation

Frank Kügler[1], Stefan Baumann[2], Antje Schweitzer[3] , Petra Wagner[4]

[1]Universität Potsdam, [2]Universität zu Köln, [3]Universität Stuttgart, [4]Universität Bielefeld
kuegler@uni-potsdam.de, stefan.baumann@uni-koeln.de

Wir präsentieren hier eine erweiterte Version der Annotationsrichtlinien von DIMA – "Deutsche Intonation – Modellierung und Annotation". DIMA ist ein Konsenssystem zur Transkription der Intonation des Deutschen, das auf den Grundannahmen des autosegmentalmetrischen Ansatzes zur Analyse von Intonation basiert (Pierrehumbert, 1980; Beckman & Pierrehumbert, 1986; Gussenhoven, 2004; Ladd, 2008) und sich dabei weitgehend auf phonetische Kriterien stützt ("*phonetically informed phonological transcription*"). Damit soll eine Vergleichbarkeit von Daten/Korpora unterschiedlicher Arbeitsgruppen im deutschsprachigen Raum gewährleistet werden. Ziel ist es, dass das System leicht erlernbar bzw. trainierbar sein soll, um eine hohe Reliabilität zwischen den Annotatoren, insbesondere zwischen Arbeitsgruppen, zu erzielen. Dies kann dadurch erreicht werden, dass die DIMAAnnotation keine abschließende phonologische Interpretation der Daten vorsieht. Diese erfolgt zu einem späteren Zeitpunkt bei der Übertragung in die jeweiligen bereits bestehenden Intonationsmodelle (vgl. z.B. Féry, 1993, 2012; Mayer, 1995; Grice, Baumann & Benzmüller, 2005; Peters, 2005, 2014). Somit stellt die Konsenstranskription eine phonetisch orientierte Repräsentation der intonatorischen Oberflächenkontur dar, die jedoch phonologische Kriterien mit einbezieht.

Das DIMA-System (Kügler et al. 2015) schlägt drei Ebenen für die prosodische Annotation vor. Diese kennzeichnen jeweils unabhängig voneinander Phrasengrenzen, Töne und die relative Prominenz von Akzenten. Auf Phrasenebene werden zwei Typen unterschieden – eine Phrase mit starker und eine mit schwacher prosodischer Grenze. Auf der Tonebene wird zwischen Akzenttönen und Nicht-Akzenttönen unterschieden. Töne sind grundsätzlich entweder hoch (H) oder tief (L) und werden relational zueinander interpretiert. Auf der Prominenzebene wird zwischen drei Stufen unterschieden: schwache, starke und extra-starke Prominenz. Die Erweiterungen der Richtlinien (Kügler & Baumann 2016) beziehen sich auf die folgenden Aspekte: (i) Zusätzliche Annotation von Disfluenzen auf der Phrasenebene, (ii) ausführlichere Kriterien für die Annotation der Tonebene, (iii) deutlichere Definition der Diakritika für die Phrasen- und Tonebene, (iv) Thematisierung möglicher Zweifelsfälle bei der Annotation.

Um die erweiterten Richtlinien zu evaluieren, stellen wir eine Studie zur Reliabilität zwischen Annotatoren vor. Hierfür haben die Autoren einen ca. zwei Minuten langen spontansprachlichen Ausschnitt eines politischen Interviews aus dem SFB 732 *silver standard collection* Korpus (Universität Stuttgart) nach DIMA annotiert. Wir berichten über die Übereinstimmung der vier Annotationen in Bezug auf die einzelnen Ebenen der Phrasierung, der Töne und der Prominenz, und diskutieren die generelle Anwendbarkeit des DIMASystems.

Referenzen

[1] Beckman, M. E., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, *3*, 255–309.

[2] Féry, C. (1993). *German Intonational Patterns*. Tübingen: Niemeyer.

[3] Féry, C. (2012). Die Annotation der deutschen Intonation: Wie kann GToBI die Syntax und die Semantik integrieren? *Linguistische Berichte*, *229*, 39–97.

[4] Grice, M., Baumann, S., & Benzmüller, R. (2005). German Intonation in Autosegmental-Metrical Phonology. In S.-A. Jun (Ed.), *Prosodic Typology.* (pp. 55–83). Oxford: OUP.

[5] Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge: CUP.
[6] Kügler, F., Baumann, S. (2016). Annotationsrichtlinien DIMA. (Version 3) MS, Universität Potsdam & Universität zu Köln.
[7] Kügler, F., Smolibocki, B., Arnold, D., Braun, B., Baumann, S., Grice, M., Jannedy, S., Michalsky, J., Niebuhr, O., Peters, J., Ritter, S. B., Röhr, C. T., Schweitzer, A., Schweitzer, K. Wagner, P. (2015). DIMA - Annotation guidelines for German intonation. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proc 18th ICPhS* (paper 317: 1-5). Glasgow, UK: University of Glasgow. Retrieved from http://www.icphs2015.info/pdfs/Papers/ICPHS0317.pdf
[8] Ladd, D. R. (2008). *Intonational Phonology* (2nd ed.). Cambridge: CUP.
[9] Mayer, J. (1995). *Transcription of German intonation: the Stuttgart System*, Ms., University of Stuttgart.
[10] Peters, J. (2005). Intonation. In *Duden-Die Grammatik* (7th ed., pp. 95–128). Mannheim: Dudenverlag.
[11] Peters, J. (2014). *Intonation*. *Kurze Einführungen in die germanistische Linguistik – KEGLI: Vol. 16*. Heidelberg: Universitätsverlag Winter.
[12] Pierrehumbert, J. B. (1980). *The phonology and phonetics of English intonation* (PhD Thesis). MIT, Cambridge, Massachusetts.

POSTER

The Karl Eberhards Corpus of spontaneously spoken southern German in dialogues – audio and articulatory recordings

Denis Arnold, Fabian Tomaschek

University of Tübingen, Germany
denis.arnold@uni-tuebingen.de, fabian.tomaschek@uni-tuebingen.de

The current paper presents a corpus containing 35 dialogues of spontaneously spoken southern German, including half an hour of articulography for 13 of the speakers. Speakers were seated in separate recording chambers, mimicking a telephone call, and recorded on individual audio channels. The corpus provides manually corrected word boundaries and automatically aligned segment boundaries. Annotations are provided in the Praat format. In addition to audio recordings, speakers filled out a detailed questionnaire, assessing among others their audio-visual consumption habits.

☞ **See full paper in Proceedings**

Acoustic correlates of word stress in German spontaneous speech

Grigorij Aronov, Antje Schweitzer

Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart
grigorij.aronov@ims.uni-stuttgart.de, antje.schweitzer@ims.uni-stuttgart.de

The acoustic properties of word stress have been explored in a number of studies. However, there is little research on German word stress, and even less on its realization in spontaneous speech. This paper tests whether parameters that have been found to implement word stress in mostly laboratory speech are also employed in a corpus of German spontaneous speech. Specifically, we consider spectral tilt, syllable duration and pitch. While the results for syllable duration conform with the prevalent finding that stressed syllables have a higher duration we find no significant effect of pitch. In the case of spectral tilt however, we observe contradicting results, depending on the way we quantify tilt.

☞ **See full paper in Proceedings**

Schwa Elision in German Utterances of Bilingual Speakers with Different Ambient Languages during Speech Acquisition

Susanne Beinrucker

Ludwig-Maximilians-Universität München
susanne.beinrucker@campus.lmu.de

The aim of this study was to investigate how a reduction phenomenon like the schwa elision is realized by bilingual speakers if it exists only in one of the two languages of the bilinguals. Therefore two simultaneous bilingual speakers of German and Hungarian, who grew up with different ambient languages, were compared. These languages suit this question because only German features the [ə] sound and its elision.
Especially it should be tested whether the ambient language – either German or Hungarian – during the speech acquisition has an influence on the schwa elision in German utterances of the bilinguals. To extend the comparison, also a monolingual native German speaker and a monolingual native Hungarian speaker, who learnt German as a foreign language, were analysed. So the prediction was that the bilingual speaker, who grew up in a German environment, would be closer to the German monolingual in the realization of schwa elision than the bilingual speaker who grew up in Hungary.
The subjects had to read aloud test sentences as an answer to a visually presented question. These recordings were analysed by listening to them and visual inspection of the sonagram. It was observed that the realizations of the bilingual, who grew up in Germany, were almost identical to the ones of the German monolingual regarding schwa elision. This confirmed the prediction. In contrast, the other bilingual realized no complete schwa elision because the schwa was still slightly perceivable but it was not really identifiable as a vowel in the sonagram. These findings offer a starting point for further research to design experiments especially with spontaneous speech and in consideration of a possible gradual schwa elision.

☞ **See full paper in Proceedings**

Deriving a strategy for synthesizing lengthening disfluencies based on spontaneous conversational speech data

Simon Betz[1;2], Jana Voße[1;3], Petra Wagner[1;2]

[1]Bielefeld University, Phonetics and Phonology Workgroup
[2]Bielefeld University, Center of Excellence Cognitive Interaction Technology (CITEC)
[3]University of Gothenburg, Department of Philosophy, Linguistics and Theory of Science
simon.betz@uni-bielefeld.de

Our overarching research project explores the usability of disfluencies in incremental spoken dialogue systems. This endeavor requires basic phonetic research on disfluencies in spontaneous speech corpora as to define strategies for synthesizing disfluencies in a meaningful way. In this paper, our current research focus lies in an investigation of disfluency-related lengthening as a promising time-buying strategy in synthesized dialogue [1][2]. We base our analyses on the results of a search tool aiming to automatically detect lengthening in spontaneous speech corpora occurring without adjacency to phrase boundaries or other disfluencies, i.e. standalone lengthening phenomena. We analyzed disfluency-related lengthening in the ”monomodal” half of the GECO corpus [3], with regard to their context, word class, syllable position and phone type. We thenpostulate a disfluency insertion strategy for synthetic speech that prioritizes lengthening phenomena based on the results obtained in our study.

☞ **See full paper in Proceedings**

A tonal analysis of the Limburgian Dialect spoken in Reuver

Sebastian Bredemann

Goethe Universität Frankfurt am Main
basti.bredemann@yahoo.de

This study aims at providing a tonal analysis of the Central Franconian dialect which is spoken in the Dutch Limburgian town of Reuver. It provides evidence for a binary lexical tone contrast, similar to that found in Roermond and Venlo, between which Reuver is geographically located. The tonal contrast, traditionally referred to as Accent I and Accent II, is only licensed in focused and final syllables. The contrast between both accents is resulting from a single lexical H-tone which is contained in accent II syllables. The interaction of this lexical tone and tones provided by intonation leads to different realizations of these accents relative to the prosodic and intonational context.
Both the dialect of Roermond and Venlo are well investigated, (for example see Gussenhoven 2004) but resemble very different patterns when it comes to the realization of Accent I and II, especially when it comes to lexical and intonational tones competing for the same prosodic edge. This study investigates in which way the dialect of Reuver differs from one or both dialects when it comes to tonal grammar, which leads to interesting insights, given that geographically Reuver is equally distant from both Roermond and Venlo.
As the theoretical framework under which the tonal contours are analyzed serves the tone-sequence model of intonation (Pierrehumbert, 1980). The tonal contours are analyzed in an Optimality theoretic framework and emphasis will be given to the arrangement of the tones from different sources (intonational and lexical) on one single tier.

☞ **See full paper in Proceedings**

Investigating the communicative function of breathing and non-breathing "silent" pauses

Aleksandra Ćwiek, Sina Neueder, Petra Wagner

Faculty of Linguistics and Literary Studies, Bielefeld University, Germany
{acwiek|sina.neueder|petra.wagner}@uni-bielefeld.de

In this study we investigate the communicative function of two types of "silent" pauses according to breathing behaviour. Taking into account the hypothesis by [1], we expected breathing pauses to be interpreted as a turn-taking cue. A question-answer study in which participants were asked to react to a question as soon as possible was conducted to test this hypothesis. Subsequent analyses of the data revealed that in comparison to nonbreathing pauses, breathing pauses are significantly more often interpreted as a turn-keeping signal, which contradicts the working hypothesis. Our results corroborate recent findings by [2].

☞ **See full paper in Proceedings**

The development of intonation and narrative abilities in a narrative retelling corpus

Conceição Cunha[1], Nicole Weidinger[3], Herta Alejo[3], Katharina Hogrefe[2], Wolfram Ziegler[2], Katrin Lindner[4], Georg Goldenberg[5]

[1]Free University of Bozen-Bolzano,
[2]Clinical Neuropsychology Research Group, LMU Munich,
[3]Institute for German as a Foreign Language, LMU Munich,
[4]Institute for German Philology, LMU Munich,
[5]Department of Neurology, Technical University Munich
cunha@phonetik.uni-muenchen.de, weidinger@daf.lmu.de,

At least in Germanic languages such as English and German, the use of pitch accents and the lack of it are crucial for the informational content of referential information. Already Halliday (1967) associated the presence or absence of a pitch accent with the distinction between new and given information, respectively. Pierrehumbert & Hirschberg (1990) added a differentiation between high and low pitch accents, with deaccentuation and low pitch marking different degrees of given information and high pitch marking new information. Bates et al. (1984) showed that intonation has already been acquired by the age of three years. However, young children generically accent new but not given information (Wieman, 1976; Gruenloh et al., 2015). For German, Müller et al. (2006) showed that monolingual 4-year-old children produced focused elements with a higher pitch than unfocused ones. Chen (2007) analysed the use of pitch accents and deaccentuation by Dutch preschool children at the age of 4-5 years. 5 and 7 years old dutch children were able to vary accentuation marking (with L+H*) new referents and deaccenting given tokens in a similar fashion as adults do (De Ruiter, 2010). However, children produced pitch accents with flatter slopes and smaller excursions than adults.

Regarding the development of narrative abilities, numerous studies revealed an expanded developmental trajectory well into primary school (e.g., Halm, 2010; Stein & Glenn, 1982): Until the age of seven/eight, children primarily produce utterances which refer to the chronological sequence of events. From the age of nine/ten, children increase the use of metanarrative structures such as comments, explanations or setting descriptions. Between the ages of six and ten, syntactic dependencies and different types of connectives marking the textual structure are used more productively by children (Colletta et al., 2010). However, the role of prosody in children's narrative production is not yet well established. This seems surprising since the importance of prosody in narrative production had been shown for adults (e.g., Hirschberg & Grosz, 1992). Studies of prosodic development mainly focussed on early language acquisition (e.g., Behrens & Gut, 2005) or examined children's use of prosody in single phrases (Wells et al., 2004). In contrast, the question of how older and more experienced children use prosodic means in spontaneous data has not attracted much interest yet (with the exception of De Ruiter (2010). At this point, this study aim to investigate developmental changes in the production of intonation of pre-schoolers and primary school's children between the age of 5 and 9, in the context of a narrative retelling task. Wells et al. (2004) suggested that intonation development continues after the age of five and that some functional contrasts are not mastered until the age of eight.

Methods. Participants were 20 five-year-old children (9 girls) between 5;1 and 5;11 (*M*=5;6) and 20 nine-year-old children (9 girls) between 9;0 and 9;11 (*M*=9;7) acquiring German. Standardized German measures of grammatical and vocabulary development, nonverbal intelligence and short term memory span confirmed children's typical development. Additionally, 20 adult native speakers

(12 women) between 21 and 58 (*M*=43) of German functioned as controls.

Material & Procedure. The stimulus material consisted of a short Silvester and Tweety cartoon (duration: 47 seconds) which included music but no words. After watching the cartoon once, the task was introduced by the neutral request "Tell me what happened in the story".

Measurements. The data was manually segmented using G-Tobi (Grice & Baumann, 2002). For the pitch range, local min and max of the fundamental frequency was measured in Hertz and converted to semitones (using a logarithm (39,863*LOG(max/min) as in Grünloh et al. (2015) for comparison reasons). Generalized linear mixed models were run for each of the variants (pitch accent types and boundary tones; pitch range).

Results. Preliminary results show an increase in the use of low pitch accents between the three groups of participants (5 < 9 < adults) and the same trend for the phrase-final rising intonation. Pitch range was significantly flatter by adults than by 5 years old children.

Discussion. This study shows that primary school children make use of a more differentiated set of prosodic possibilities than preschoolers, since they differentiate H* and L* pitch accents and signal continuation of the story using raising boundary tones.

References.

[1] Arnold, J. (2008). "THE BACON" Not "the Bacon": How Children and Adults Understand Accented and Unaccented Noun Phrases, Cognition, 108, 69-99.

[2] Bates, E., MacWhinney, B., Caselli, C., Devescovi A., Natale, F., & Venza, V.(1984). A cross-linguistic study of the development of sentence interpretation strategies. Child Development 55, 341-354.

[3] Behrens, H. & Gut, U. (2005). The relationship between prosodic and syntactic organization in early multiword speech. Journal of Child Language, 32, 1-34.

[4] Chen, A. (2007). Intonational realisation of topic and focus by Dutch-acquiring 4- to 5-year-olds. In J. Trouvain, & W. J. Barry (Eds.), Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS 2007), 1553-1556.

[5] Colletta, J.-M., Pellenq, C., & Guidetti, M. (2010). Age-related changes in co-speech gesture and narrative: Evidence from French children and adults. Speech Communication, 52, 565-576.

[6] Grice, M. & Baumann, S. (2002). Deutsche Intonation und GToBI. Linguistische Berichte 191, 267-298.

[7] Grünloh, T., Lieven, E., & Tomasello. M (2015). Young Children's Intonational Marking of New, Given and Contrastive Referents. Language Learning and Development. 95-127.

[8] Halliday, M. (1967). Notes on Transitivity and Theme in English, Journal of Linguistics, 3, 199-244.

[9] Halm, U. (2010). Die Entwicklung narrativer Kompetenz bei Kindern zwischen 7 und 14 Jahren. Marburg: Tectum Verlag.

[10] Hirschberg, J. & Grosz, B. J. (1992). Intonational features of local and global discourse structure. Proceedings of the workshop on speech and natural language, 441-446.

[11] Müller, A., Höhle, B., Schmitz, M., & Weissenborn, J. (2009). Information structural constraints on children's early language production: The acquisition of the focus particle auch ('also') in German-learning 12- to 36-month-olds. First Language, 29(4), 373-399.

[12] Pierrehumbert, J. B. & Hirschberg, J. (1990). The Meaning of Intonational Contours in the Interpretation of Discourse. In: P. R. Cohen, J. Morgan, M. E. Pollack, (eds.), Intentions in Communication. Cambridge: MIT Press. 271-311.

[13] de Ruiter, L. (2010). Studies on intonation and information structure in child and adult german. PhD thesis, Max Planck Institute for Psycholinguistics, Nijmegen.

[14] Stein, N. & Glenn, C. (1979). An analysis of story comprehension in elementary school children. In: Freedle, R. O. (Ed.), Advances in discourse processing (Vol. 2). Norwood, NJ: Ablex.

[15] Wells, B., Peppe, S. & Goulandris, N. (2004). Intonation development from five to thirteen, Journal of Child Language, 31, 749-778.

[16] Wiemann, L. A. (1976): Stress patterns in early child language. Journal of Child Language, 3, 283-286

PresenterPro: A tool for recording, indexing and processing prompted speech with Praat

Volker Dellwo

Department of Computational Linguistics, University of Zurich
volker.dellwo@uzh.ch

Praat (www.praat.org) is a powerful tool for a wide variety of speech analysis and processing tasks. When it comes to recording speech, however, it lacks some fundamental functions that allow a user to prompt a reader with a written list of words or sentences (henceforth: speech prompts) on a screen and index the prompted recordings for further processing. *PresenterPro* - a Praat plug-in - fills this gap. It (a) prompts a reader to read utterances from a screen, (b) automatically indexes the recorded speech prompts in a Praat TextGrid and (c) extracts all recorded speech prompts into individual files. It thus offers an efficient solution for recording large lists of speech prompts. The present paper describes the plug-in and discusses in which situations it is particularly useful.

☞ **See full paper in Proceedings**

Assessment of Prosodic Attributes in Codec-Compressed Speech

Johanna Dobbriner, Oliver Jokisch, Michael Maruschke

Hochschule für Telekommunikation Leipzig, Institut für Kommunikationstechnik, Germany
{johanna.dobbriner|jokisch|maruschke}@hft-leipzig.de

This article deals with the representation of prosodic attributes in coded speech which is less-studied. Common models in speech coding assume that there is no relevant influence of prosodic variation on perceived quality and content of coded speech under suitable operating conditions. Our experiments included a listening test and the instrumental assessment of utterances from an especially constructed test database for the three categories *focus*, *type of sentence* and *situation*. Each category contained at least three different text phrases in several variants, and each original sample was compressed using the fullband-audio Opus codec and the narrowband G.711 codec for reference. The listeners evaluated the overall speech quality and processed a matching task to given prosodic categories. In general, the prosodic variations were well-recognised even when the coding degradation was significant. The overall assessments were comparably high, by achieving an MOS of 4.3 and above on the five-point scale. The hybrid Opus coding method seems to maintain the prosodic features of speech as given in the original reference.

☞ **See full paper in Proceedings**

Wahrnehmungsexperimente mit Hilfe eines Computerspiels

Daniel Duran, Natalie Lewandowski, Antje Schweitzer

Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, Deutschland
daniel.duran@ims.uni-stuttgart.de

Wir präsentieren ein neu entwickeltes Computerspiel zur Durchführung von Sprachwahrnehmungsexperimenten. Klassische computergestützte Experimente in der Sprachwahrnehmungsforschung stellen keine natürliche Versuchsumgebung dar. Sie sind hochgradig überwacht, lenken die Aufmerksamkeit der Versuchsteilnehmer explizit auf bestimmte zu untersuchende Aspekte und sie unterscheiden sich deutlich von natürlichen Gesprächssituationen. Eine Lösung für dieses Problem ist der Einsatz eines Computerspiels, bei dem Aufmerksamkeit auf phonetische Details sich aus der Spielsituation ergibt. Computerspiele finden immer weitere Verbreitung in der Psychologie oder sprachwissenschaftlichen Verhaltensexperimenten. Unsere neue Experimentumgebung implementiert ein klassisches Kategorisierungsexperiment in Form eines Computerspiels. Mit einer modernen Spiel-Engine wurde ein Egoshooter entwickelt, in welchem die Spieler so schnell wie möglich bestimmte Objekte auf dem Bildschirm anklicken müssen. Die Spieler müssen in einer dreidimensionalen Umgebung auf Stimuli reagieren, die durch Spielfiguren dargestellt werden. Die zwei Antwortkategorien werden zunächst durch akustische sowie visuelle Merkmale dargestellt. Die Unterscheidung ist dann zunehmend nur noch anhand der akustischen Stimuli möglich. Die Spieler werden durch diese Spielumgebung motiviert das zu Grunde liegende Kategorisierungsexperiment möglichst gut zu lösen ohne dabei in einer völlig unnatürlichen Situation explizit darauf hingewiesen zu werden. Wir diskutieren praktische sowie theoretische Aspekte unseres Spiels und präsentieren erste Erfahrungen damit aus einer Perzeptionsstudie mit manipulierten phonetischen Details in natürlichen Sprachstimuli.

☞ **See full paper in Proceedings**

The temporal relation between talker and word recognition

Michael Fischbacher, Eva Reinisch

Institute of Phonetics and Speech Processing, Ludwig Maximilians University Munich
evarei@phonetik.uni-muenchen.de

The speech signal not only contains information about the meaning of what is said, but also about the talker who produces it. Since talker information is not necessary to understand an utterance it has been suggested that it is discarded during speech processing (e.g. [1]). However, recently, a growing body of research has demonstrated an interrelation between talker and word recognition (e.g. [2], [3]).
The aim of the present study was to shed light on the nature and timing of this interrelation, since previous results as to whether the use of speaker information helps or harms word recognition are mixed. In addition, although there is evidence that word information is processed faster than talker information, the suggested times vary from study to study.
As a method, we adapted the visual-world eye-tracking paradigm, which is based on the effect that listeners spontaneously fixate visual referents (i.e., pictures) of spoken input in a closely timelocked fashion. This allows tracking of the on-line processing of phonetic information during word recognition (e.g. [4]). For our study we complemented typical pictures of word referents with pictures of the talkers. This allowed us to simultaneously record the on-line processing of talker information (see [5] for a pilot study on this method).
In a three-part experiment, thirty-two participants were asked to first listen to words and click on the picture matching the word. Words were produced by two male and two female speakers who were unknown to the participants. Next, participants learned to match the four voices to pictures of the talkers. Finally, in the main part of the experiment, participants were asked to click on the matching referent while viewing combinations of the referent pictures and the talkers. The critical question was whether in this last part, participants would fixate more on the correct word referent or the talker, and in what temporal relation.
Results confirmed an interrelation between talker and word recognition as both types of referents were fixated on, suggesting that both types of information were activated in on-line speech processing. A comparison of the first and third part of the experiment suggests that talker recognition supported word recognition, since the target was fixated more and earlier when pictures of the talkers were shown in addition to the word referents. Moreover, participants who preferred clicking on talker over the word pictures solved the task fastest. Since words were acoustically ambiguous up to their third phoneme, information about who is speaking could have started to modulate word recognition and influence target selection.
In summary, our results suggest that talker information is all but discarded during word recognition. In line with a number of previous studies, we could demonstrate that talker information can positively influence spoken-word recognition. By adapting the visual-world eye-tracking paradigm to track more than simple word recognition, we opened a window into tracking the complex relations between different types of information that listeners might potentially exploit in order to efficiently understand what is being said. Issues about the temporal uptake of different types of information will be discussed.

References

[1] Liberman, A. & Mattingly , I. (1985). The motor theory of speech perception revised. Cognition, 21, 1–36.
[2] Cutler, A., Andics, A., Fang, Z. (2011): Inter-dependent categorization of voices and segments. ICPhS Hong Kong.
[3] Creel, S. & Tumlin, M. (2011). On-line acoustic and semantic interpretation of talker information. Journal of Memory and Language, 65, 264–285.
[4] Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. Journal of Memory and Language, 38, 419–439.
[5] Schindler, C. & Reinisch E. (2015): Tracking the temporal relation between speaker recognition and processing of phonetic information. Proceedings of the 18th International Congress of Phonetic Sciences. Glasgow.

Die akustischen und artikulatorischen Korrelate des /r/ im Norddeutschen. Eine Ultraschallstudie.

Riccarda Funk, Christina Otto

Institut für germanistische Sprachwissenschaft, Friedrich-Schiller-Universität Jena
riccarda.funk@uni-jena.de, otto.christina@uni-jena.de

Die hier vorgestellte Studie untersucht, wie der /r/-Laut in der standardnahen norddeutschen Varietät im Silbenanlaut und im Silbenauslaut nach Langvokal artikulatorisch und akustisch realisiert wird. Im Silbenonset wurden Minimalpaare mit /r/ oder /h/ und in der Silbencoda Wörter mit /r/ nach Langvokal in Trägersätze eingebettet. Diese wurden artikulatorisch durch Ultrasound Tongue Imaging mittels *AAA* [1] und akustisch (F1, F2) von 10 Sprechern aus dem Norddeutschen Raum untersucht.
Die Analyse zeigt die Bildung einer dorso-uvularen Enge bei /r/ im Silbenonset im Vergleich zu /h/. Bei /r/ nach Langvokal kommt es zu einer Diphthongbildung, deren Ausprägung in Abhängigkeit zum Vokal auftritt. Je näher sich der Ausgangsvokal an [ɐ] befindet, umso schwächer ausgeprägt ist die Zungenbewegung.

☞ **See full paper in Proceedings**

A Shadowing Experiment with Natural and Synthetic Stimuli

Iona Gessinger[1], Eran Raveh[1], Johannah O'Mahony[1]
Ingmar Steiner[1,2], Bernd Möbius[1]

[1]Saarland University,
[2]DFKI GmbH
lastname@coli.uni-saarland.de

Inter-speaker accommodation is a phenomenon observed in human communication. Phonetic convergence is one way for a speaker to accommodate to an interlocutor. It is defined as an increase in segmental and suprasegmental similarities between two speakers [1]. Phonetic convergence has been found for human-to-human interaction in both spontaneous, conversational speech [1, 2] and non-conversational speech occurring in experimental settings such as the shadowing task [3, 4]. Previous studies on convergence in human-to-human interaction looked at suprasegmental features such as f0 range [5] and speaking rate [6], as well as segmental features such as spectral properties of vowels [3] and voice onset time [7]. Thus far, phonetic convergence has received little to no attention in the field of human-computer interaction (HCI). In the experiment introduced in this paper, we take a first step in investigating whether human speakers also converge to synthesized speech by conducting a shadowing experiment using both natural and computergenerated stimuli, concentrating on selected segmental
features (cf. 2.1).
Based on previous findings in human-human interaction, we expect to observe phonetic convergence on the segmental level for the natural stimuli. Since the quality of synthesized speech is improving and HCI is becoming ever more used for various tasks in everyday life, humans are likely to interact in a similar way with computers as they do with humans. Therefore, we expect to observe convergence for the synthetic stimuli as well. However, the degree of convergence might still be influenced by the perceived naturalness of the synthetic stimuli.

☞ **See full paper in Proceedings**

Wenn Stotterer nicht stottern. Quantifizierung dynamischer Ultraschalldaten

Cornelia J. Heyde [1], *James M. Scobbie* [1]

Clinical Audiology, Speech and Language (CASL) Research Centre, Queen Margaret University, Edinburgh, UK
cheyde@qmu.ac.uk, jscobbie@qmu.ac.uk

Stottern wird traditionell akustisch definiert, wobei man sich oftmals ausschließlich auf akustisch-perzeptive Unterbrechungen im ansonsten flüssigen Redefluss stützt. Die detaillierte artikulatorische Analyse der flüssigen Sprache von Stotterern soll Auskunft darüber geben, ob Elemente des Stotterns eventuell selbst dann nachweisbar sind, wenn sie akustisch nicht wahrnehmbar sind. Eine weitere Frage, die wir mit der dyna-mischen Analyse der Ultraschalldaten beantworten wollen, ist, wo genau sich Stottern manifestiert. Mit Hinblick auf die Fault-Line Hypothese von Wingate [1] untersuchen wir insbesondere die Bewegung der Zunge in die Verschlußstellung (Anglitt) und vergleichen diese mit der Bewegung im Über-gang von der Verschlußstellung zum darauffolgenden Vokal (Abglitt). Die Ergebnisse unserer Untersuchung deuten an, dass sich Stotterer selbst in der scheinbar flüssigen Sprache von Kontrollsprechern unterscheiden. Artikulatorische Unterschiede zwischen den beiden Sprechergruppen wurden im Übergang von Konsonant zu Vokal beobachtet, was die Annahme Wingate's bestärkt, dass beim Stottern das Problem nicht auf einem bestimmten Wort oder Laut, sondern im Übergang von einem zum nächsten Segment, liegt.

☞ **See full paper in Proceedings**

Categorial and acoustic changes to rhoticity in bilinguals

Marie-Christin Himmel, Barış Kabak

Universität Würzburg
marie-christin.himmel@uni-wuerzburg.de, baris.kabak@uni-wuerzburg.de

Rhotic sounds are subject to a great deal of cross-language variation. For instance, synchronic as well as diachronic observations of English postvocalic /r/ show that it is subject to variable degrees of realization (full, semi- and non-rhoticity). Furthermore, bilinguals with long-term exposure to a rhotic second language (L2) have been shown to differ in their acoustic realizations of segmental and suprasegmental patterns in their native language (L1, e.g. Flege & Efting 1987, Mennen 2004), including the acoustic properties of [Vr]-sequences (cf. Ulbrich & Ordin 2014). Additionally, variable realizations of /r/ in semi-rhotic English varieties suggest that the loss/emergence of the postvocalic /r/ both within and across life spans may be based on an inherent vulnerability of acoustic patterns found in rhotic sequences, which have been previously described as fragile (cf. Lutz 1994).

Here, we shed light on rhotics from the perspective of adjustments in L1 due to long-term exposure to an L2 and investigate whether adjustments to postvocalic /r/ within a life span follow universal principles as well as diachronic trajectories. Furthermore, whether non-rhotic L2s may exert an influence on L1 rhoticity (i.e., the opposite direction described in Ulbrich & Ordin 2014) has hitherto not been explored.

In our experimental study, we investigated changes in the phonetic quality of /r/ in the native language of American English speakers (from rhotic varieties), who had post-pubescent long-term exposure to L2 German (a non-rhotic language), and tested which phonological contexts variables favor the deletion of postvocalic /r/.

Twelve American English-German late bilinguals residing in Germany (mean LOR=25 years) performed a variety of speech elicitation tasks in both their L1 and L2. Additionally, a matched monolingual control group (n=9) was recruited for the test battery. The test items used in our elicitation tasks were controlled for preceding vowel as well as syllable and morphological complexity. Auditory coding into binary categories (rhotic/non-rhotic) was used to analyze presence and absence of rhoticity. Subsequent acoustic analysis was performed on the tokens identified as rhotic using Praat (Boersma & Weenink, 2016). Measurements were taken semi-automatically at (i) the peak of F3 within the vowel portion and (ii) the lowest point of F3 within the /r/-portion of the [Vr]-sequence. Following Hagiwara (1995), we also employed a neutral F3 value by measuring the F3 of plain vowels as a reference point for each participant to normalize raw Hertz values and used percent ranks of the F3 slope within the [Vr]-sequence.

Our results show differences both in the categorical as well as in the acoustic analysis: The categorical data revealed that, as compared to monolinguals, the realization of /r/ by bilinguals exhibited a high degree of variability, albeit constrained by various factors such as quality of the preceding vowel, syllable position, and task. Importantly, the pattern of change here mirrors observations from studies on both synchronic and diachronic variation of rhoticity in English.

Our acoustic analyses of the rhotic [Vr]-sequences show a significant effect of group for F3 within the vowel portion, indicating that bilinguals produce a lower degree of r-coloring in the vowel section preceding a postvocalic /r/. Additionally, we find a significant effect of group for F3 in the /r/-portion of the [Vr]-sequence, indicating that bilinguals also show changes in their realization of the postvocalic /r/ itself. Here, we did not find a significant interaction between group and other

factors. In the vowel proportion of the [Vr]-sequences, a significant interaction for group and task indicates that acoustic variation is conditioned by the level of formality and possibly by the type of interlocutor, whereas bilinguals show no sensitivity to such variables.
We take our results to show that non-rhoticity in the L2 influences rhoticity in the L1 in terms of (i) the gradual loss of post-vocalic /r/, as well as (ii) the acoustic quality of [Vr]-sequences. We will suggest that acoustic changes in rhoticity appear across the board, whereas absence and presence of rhoticity within an individual's lifespan mirrors both diachronic and synchronic patterns commonly observed in English varieties and thus is subject to language-specific constraints.

References

[1] Boersma, P. & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.19, retrieved from http://www.praat.org/

[2] Flege, J.E. & Eefting, W. (1987). The production and perception of English stops by Spanish speakers of English. *Journal of Phonetics* 15: 67–83.

[3] Hagiwara, R. (1995). *Acoustic Realizations of American /r/ as Produced by Women and Men* (PhD dissertation). University of California, Los Angeles, USA.

[4] Lutz, A. (1994). Vocalization of "post-vocalic *r*": An Early Modern English sound change? In D. Kastovsky (ed.), *Studies in Early Modern English* (pp. 167-185). Berlin: Mouton de Gruyter.

[5] Mennen, I. (2004). Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics* 32: 543–563.

[6] Ulbrich, C. & Ordin, M. (2014). Can L2-English influence L1-German? The case of post-vocalic /r/. *Journal of Phonetics* 45: 26-42.

The realisation of Albanian laterals in German as a second language: A case study

Bettina Hobel, Sylvia Moosmüller, Christian Kaseß

Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria
{bettina.hobel|sylvia.moosmueller|christian.kasess}@oeaw.ac.at

In this case study, it is investigated how laterals spoken by one female Albanian speaker are realised in German as a second language. Standard Albanian features two lateral phonemes, an alveolar and a velarised one, whereas Standard Austrian German (SAG) has only one phoneme but a variety of possible realisations ranging from alveolar laterals in the standard language to velarised laterals, retroflex laterals, and l-vocalisations in the dialects. The results of our Albanian speaker show that she mostly produces alveolar laterals, occasionally with a tendency towards palatalization. Only two velarised laterals occurred, both in position between consonants. Hence, it can be concluded that the Albanian speaker abandoned one of the two Albanian lateral phonemes, namely the velarised lateral, and instead, she only uses the alveolar lateral phoneme of SAG. Concerning the l-vocalisation and retroflex laterals, as found in the Styrian dialects, only a few instances occurred; in general, the realisation of l-vocalisation was restricted to unstressed positions. The phonetic context has significant influence on the F2 of the lateral, which mainly involves differences between previous and following front vowels versus consonants. F2 of laterals preceded or followed by front vowels is higher than preceded or followed by consonants.

☞ **See full paper in Proceedings**

Fully Automated Accent Correction for Computer-Assisted Speech Rhythm Training: A Pilot Study

Markus Jochim, Christoph Draxler

Institute of Phonetics and Speech Processing, Ludwig Maximilian University Munich
{markusjochim|draxler}@phonetik.uni-muenchen.de

In this project we have implemented a feedback method for CAPT (computer-assisted pronunciation training), where the learner's syllable durations are corrected in a fully automated way. The correction is based on the PSOLA algorithm [1] and uses a reference speaker's recording to determine the desired syllable durations for a given utterance. Learners get to hear a pre-defined sentence in their target language and are expected to imitate it as closely as possible. Their recording is then manipulated to match the reference speaker's syllable durations and the manipulation is presented as auditory, corrective feedback. Exploiting this feedback, learners are expected to try again and improve their pronunciation in terms of speech rhythm. A pilot study with six learners of German with various L1 backgrounds was conducted to assess whether learners can successfully exploit the feedback to improve their pronunciation. The results suggest that learners can improve their pronunciation significantly using this method. They also suggest that the method works equally well or better than repeatedly listening to the reference speaker.

☞ **See full paper in Proceedings**

Speech in Interaction – The Zurich Tangram Corpus

Yshai Kalmanovitch

University of Zurich, Switzerland
Yshai.kalmanovitch@uzh.ch

The Zurich Tangram Corpus (ZTC) was design primarily to investigate the relationship between interactional experience and inter-speaker phonetic convergence. It is aimed to look primarily at phonetic convergence between interlocutors as a factor of interactional intensity and of mutual interactional experience over time. It contains recordings of dyad interaction in semi-natural discourse. Subjects were required to work together on an order reconstruction task in three sessions one week apart from each other. In addition to audio data, video data as well as eye-tracking data were collected.

☞ **See full paper in Proceedings**

Der optionale Komplementierer im Deutschen – ein Fall prosodischer Syntax

Gerrit Kentner1, Isabelle Franz[1], Christian Dück[1]

[1]Goethe-Universität Frankfurt
kentner@lingua.uni-frankfurt.de, i.franz@em.uni-frankfurt.de, chrisd14@gmx.net

Wir untersuchen das Vorkommen von eingebetteten Verbzweit und Verbend-Sätzen (letztere mit dem Komplementierer “dass”) in Abhängigkeit der rhythmisch-prosodischen Struktur der Satzspitze des eingebetteten Satzes. Die Korpusdaten legen einen deutlichen Einfluss des linguistischen Rhythmus auf die Syntax nahe. Bemerkenswert ist dabei, dass der Rhythmus nicht nur die Wortabfolge beeinflusst, sondern bereits die Wahl der syntaktischen Struktur (Haupt- vs. Nebensatz) bedingt. Für Modelle der Sprachproduktion bedeutet dieses Ergebnis, dass die Prosodie nicht nur die syntaktische Struktur reflektiert, sondern die Wahl der syntaktischen Struktur mitbestimmt.

☞ **See full paper in Proceedings**

New evidence for prosodic parallelism in German(ic) morphophonology

Gerrit Kentner

Goethe-Universität Frankfurt
kentner@lingua.uni-frankfurt.de

This paper presents two studies that make the case for prosodic parallelism as a factor in German(ic) word formation.

☞ **See full paper in Proceedings**

Relation between articulatory and acoustic information in phonemic representations

Eugen Klein[1], *Jana Brunner*[1], *Phil Hoole*[2]

[1] Humboldt University of Berlin
[2] Ludwig Maximilian University of Munich
eugen.klein@hu-berlin.de, jana.brunner@hu-berlin.de, hoole@phonetik.uni-muenchen.de

The Direct Realism approach to speech perception [5] assumes that the acoustic signal produced by a speaker is interpreted by the listener in terms of articulation. This instantaneous mapping between acoustic properties of a speech sound and the articulatory configuration which produced it, assumes a strong causal relation between the two. This relation was investigated here by prompting native Russian speakers to produce the sound /ɨ/ by means of two distinct articulatory configurations resembling Russian steadystate vowels /i/ and /u/. The dissociation between articulation and acoustics is achieved through systematic changes in the direction of auditory perturbation of the second formant (F2) produced by the participants. We report first results of a study which extends empirical knowledge about the role of articulatory information in speakers' phonemic representations.

☞ **See full paper in Proceedings**

Der Einfluss der F0-Kontur als akustischer cue für chinesische Deutschlerner in konkurrierenden Kontexten

Nicola Klingler

Institut für Schallforschung – Österreichische Akademie der Wissenschaften
nicola.klingler@gmail.com

Der lexikalische Ton prägt die chinesischen Sprachen nicht nur intralingual, sondern stellt für chinesische Sprecher darüber hinaus das salienteste Merkmal zur perzeptiven Identifikation von Silben auch außerhalb der chinesischen Sprachen dar (siehe bspw. [1] für das Englische). Dass dabei der sprachspezifische Gebrauch der F0-Kontur in der target language (TL) eher unwichtig ist, zeigen die Ergebnisse des durchgeführten Experiments, welches die Perzeption des deutschen Wortakzent untersuchte: Obwohl die F0-Kontur im Deutschen meist nur zur Markierung der Intonation – und somit zur pragmatischen Differenzierung – oder zur Markierung des Fokusakzents in den Vordergrund tritt [2], vertrauten die in diesem Perzeptionsexperiment untersuchten chinesischen Sprecher so stark auf den cue F0-Kontur, dass dieser, auch wenn er in Konkurrenz mit divergierenden cues (hier: Vokallänge, Vokalqualität und Intensität) stand, einen marginal signifikanten Einfluss auf die Wahrnehmung hatte.

☞ **See full paper in Proceedings**

The interplay of prominence and boundary strength: a comparative study

Katalin Mády[1], Felicitas Kleber[2], Uwe D. Reichel[1], Ádám Szalontai[1]

[1]Research Institute for Linguistics, Hungarian Academy of Sciences, Hungary
[2]Ludwig-Maximilians-Universität Munich, Germany
{mady.katalin|uwe.reichel|szalontai.adam}@nytud.mta.hu, kleber@phonetik.uni-muenchen.de

Hungarian is a language with left-headed head-/edgeprominence. Our goal was to investigate if prominence in Hungarian can be increased by inserting or strengthening phrase boundaries before emphasised words. German, being a rightheaded head-prominence language was the basis for the comparison. Since prominence marking in Hungarian is highly dependent on syntax, a list of fruits differing in size was used. Participants were asked to utter fruit names so that someone else can guess if a fruit was small or large. We hypothesised that Hungarian speakers would use boundary signals preceding a large fruit, whereas Germans would either insert boundaries after a large fruit or not make use of final lengthening and pauses. Results show that Hungarians use more pauses than Germans in all positions, and the occurrence of pauses is used to enhance prominence. While pre-boundary lengthening was only observed preceding a large fruit in Hungarian, it was present for speakers of both languages in the final syllable of the large fruit itself. Pause occurrences after a large fruit did not depend on fruit size in any of the languages.

☞ **See full paper in Proceedings**

How to distinguish between self- and other-directed wh-questions?

Katalin Mády, Uwe D. Reichel

Research Institute for Linguistics, Hungarian Academy of Sciences, Hungary
{mady|uwe.reichel}@nytud.mta.hu

The most general aim of wh-questions is to seek for information, but they can have a wide range of other pragmatic functions. In this paper we investigate self-directed questions in dialogues that are lexicalised forms of vacillation ("how should I explain?") and do not directly address the interlocutor. Their prosodic properties are compared with real wh-questions that seek for information.

☞ **See full paper in Proceedings**

Perception of Pitch Scaling in Rising Intonation on the Relevance of f0 Median and Speaking Rate in German

Jan Michalsky

Institute of German Studies, University of Oldenburg, Germany
j.michalsky@uni-oldenburg.de

Although phonetic in nature, pitch scaling is described to as-sume certain linguistic functions such as marking of focus or sentence mode [1, 2, 3, 4, 5]. Accordingly, it is linguistically relevant to understand how precisely pitch scaling is perceived. Recent evidence suggests that the excursion size of f0 in semitones might not be a stable cue across speakers and registers [5]. As an alternative, De Looze and Hirst [6] propose to relate f0 measurements to the speaker's median. Two perception experiments on the evaluation of scaling of final rises in German were conducted to investigate which measurement is best suited to represent the perception of f0 movements. The results indicate that the distance between a specific f0 meas-urement and the speaker's median in semitones is the most stable parameter in describing pitch perception within speakers. On the other hand, no measurement solely based on f0 succeeded in explaining the observed differences in perception across speakers. A post-hoc analysis showed that speaking rate might be a non-intonational feature, which influences expectations about a given speaker's natural pitch range. Accordingly, we propose a measurement that incorporates speaking rate to adequately describe f0 movements with respect to pitch perception.

☞ **See full paper in Proceedings**

Effects of perceived attractiveness and likability on global aspects of fundamental frequency

Jan Michalsky, Heike Schoormann

Institute of German Studies, University of Oldenburg, Germany
j.michalsky@uol.de, heike.schoormann@uol.de

Acoustic parameters of the speech signal such as overall mean pitch, pitch range or variability of pitch can influence how listeners evaluate attractiveness as well as likability. This study asks two questions: Firstly, does the perception of the interlocutor's attractiveness or likability in turn influence the speaker's pitch behavior as well? Secondly, are those potential effects influenced by the interlocutor's pitch behavior in terms of entrainment? We conducted a speed dating experiment with 20 speakers in 100 mixed-sex pairs and analyzed acoustic correlates of the speakers' pitch as well as their evaluations of the interlocutors' attractiveness and likability. For both sexes, the results show a positive correlation of the speakers' pitch range with perceived attractiveness of the interlocutor and a positive correlation of the overall mean with the degree of perceived likability. Additionally, speakers showed a relative adaptation to the interlocutors' pitch which strengthened the effects of likability but diminished the effects of attractiveness. Conclusively, we suggest that speakers' pitch features are influenced by their perception of the interlocutor. However, adaptation to the interlocutors' pitch strongly interferes with these effects making it imperative to control for entrainment when investigating pitch effects of social variables.

Effects of perceived attractiveness and likabilityon global aspects of fundamental frequency☞
See full paper in Proceedings

Opa vs Oper: Neutralization of /ɐ/ and unstressed /a/ contrast in a perception and production study.

Bokelmann, J.[1]; Ellert, M.-T. [1]; Friesenhan, N. [1]; Generlich, G. [1]; Hiebert, Y. [1]; Malon, T. [1]; Naumann, F. [1]; Sachs, R. [1]; Swiech, A. [1]; Walak, M. [1]; Yovcheva, M. [1]; Zander, I. [1]; Rathcke, T.[2]; Mooshammer, C. [1]

[1] Institut für deutsche Sprache und Linguistik, Humboldt Universität zu Berlin, Germany
[2] English Language and Linguistics, University of Kent, UK
T.V.Rathcke@kent.ac.uk, christine.mooshammer@hu-berlin.de

The present study examined differences in production and perception of the German vowels /a/ and /ɐ/ in word-final, unstressed position. In the first experiment, 3 male and 3 female speakers produced minimal pairs embedded in meaningful sentences and varied in prosodic environment. In the second experiment, the minimal pairs were extracted from the context and presented to 44 listeners for a forced-choice identification task. Results showed a better-than-chance performance that was, however, mainly driven by one male speaker. Temporal and spectral measures confirmed that only this speaker produced an acoustic difference between /a/ and /ɐ/.

☞ **See full paper in Proceedings**

Universal phonetics revisited: Eine cross-linguistische Untersuchung zum Einfluss der Stimmhaftigkeit des Folgekonsonanten auf die Vokallänge im Polnischen und Deutschen

Katharina Nimz[1], Judith Baumann[1], Arkadiusz Rojczyk[2]

[1]Fakultät für Linguistik und Literaturwissenschaft, Universität Bielefeld
[2]Institute of English, University of Silesia
katharina.nimz@uni-bielefeld.de, judith.baumann@web.de, arkadiusz.rojczyk@us.edu.pl

Die vorliegende Studie umfasst zwei Produktionsexperimente, die sich mit dem Einfluss der Stimmhaftigkeit des Folgekonsonanten auf die Vokallänge im Deutschen und Polnischen befassen. Beide Sprachen sind prominente Beispiele für die phonologische Regel der Auslautverhärtung, weshalb die Mehrzahl der vorherigen Studien Vokallänge im Rahmen von *Incomplete Neutralization* untersucht hat. Bisher existieren kaum Daten zu Vokallängen im Deutschen und Polnischen in nicht-neutralisierenden Kontexten; unsere Studie greift dieses Forschungsdesiderat auf.

☞ **See full paper in Proceedings**

Wer die Qual hat, hat keinen Wal: Orthographische Effekte bei der Produktion deutscher Vokale

Katharina Nimz, Katharina Immel, Kai Ole Koop

Fakultät für Linguistik und Literaturwissenschaft, Universität Bielefeld
{katharina.nimz|katharina.immel|kkoop}@uni-bielefeld.de

In dem vorliegenden Pilotprojekt wurde untersucht, ob und inwiefern sich die explizite orthographische Markierung von Vokallänge (z. B. Dehnungs-*h*) auf die Produktion deutscher Langvokale ausübt. In einem Produktionsexperiment mit neun deutschen MuttersprachlerInnen wurden mittels Definitionen und Lückensätzen (ohne orthographischen Input) die Produktionen von 16 heterographen Minimalpaare erhoben (z. B. *Wahl* und *Wal*), deren Vokallänge mittels PRAAT ermittelt wurde. Obgleich diese Paare als homophon gelten, zeigte sich ein signifikanter Längenunterschied zwischen explizit markierten und unmarkierten Testitems: Vokale werden ca. 8% länger artikuliert, wenn sie orthographisch markiert sind.

☞ **See full paper in Proceedings**

Vocal Fry – A Marker of Sophistication or Stupidity?

Amra Odobasic

Ruprecht-Karls-Universität Heidelberg
amra.odobasic@alumni.uni-heidelberg.de

The purpose of this paper is to give an overview of vocal fry and discuss the ambiguity regarding its evaluation. Formerly classified as part of a clinical voice disorder, vocal fry (or creaky voice, pulse register, glottalization, etc.) appears to be omnipresent in American English today.
Its acoustic characteristics include a combination of rapid and short glottal pulses, as well as a very low fundamental frequency. The sporadic use of vocal fry has been found to have structural and pragmatic purposes. It is commonly used as a marker of phrase and sentence boundaries and, in RP, indicates the end of a speaker's turn in conversation.
Sociolinguistically, it might also be a cue to speaker identification and a gender marker. Studies show that even though young men have begun speaking in vocal fry more frequently, a vast quantity of young women already use it on a regular basis. This fact has been cause to rising polemics in newspaper articles and online videos: Vocal fry has been referred to as an "epidemic", making the female speaker appear uneducated, annoying, and shallow.
While the mainstream opinion seems to be rather negative, there are studies that have shown that vocal fry is also evaluated as a sophisticated trait, being associated with high social status and authority and making the female speaker seem "professional", "urban" and "upwardly mobile".

☞ **See full paper in Proceedings**

Musicality and Auditory Acuity

Miriam Oschkinat [1,2], Jana Brunner [2], Phil Hoole [1]

[1] Ludwig-Maximilians-University Munich,
[2] Humboldt-University Berlin
Miriam.Oschkinat@gmx.net, Jana.Brunner@hu-berlin.de, Hoole@phonetik.uni-muenchen.de

The perception of speech is a major field of interest in phonetic research. In most psychoacoustic perception experiments subjects are controlled with respect to age, gender, education and more. One feature which is mostly not recorded is musical education. Especially in perception tasks, however, musicality should be controlled, since musical education is linked to trained listening skills, which might be an advantage not only for the perception of music but also for the perception of speech.

Particularly for experiments treating perception of pitch, intonation and prosody musicality seems to be influential (Pape, 2009). Kishon-Rabin, Amir, Vexler et al. (2001) show that musicians identify smaller changes in frequency than non-musicians. In addition, the authors investigated a difference between classically and contemporarily educated musicians and showed that classically educated musicians perform better than contemporarily educated musicians. Also, musicians with a longer duration of musical practice (in years) performed better than the others. Another interesting aspect concerning some musicians is the ability to perceive absolute pitch, which is associated with neurophysiological adaptions and morphological changes in the *planum temporale*, a brain area which is also involved in auditory and speech perception processes (Oechslin, Meyer, Jaencke, 2010). Masataka (2011) tested absolute pitch possessors (AP) and relative pitch possessors (all musicians without AP) on speech relevant auditory acuity and showed that absolute pitch possessors were faster in identifying isolated syllables than relative pitch possessors.

The present study investigates a possible impact of musical education and absolute pitch possession on auditory acuity. We expect musicians to perform better in auditory acuity tasks than non-musicians, due to their trained listening. Furthermore, absolute pitch possessors are expected to show an even higher performance, especially in vowel discrimination tasks, as a consequence of their ability to identify subtle changes in frequency in a more precise manner than relative pitch possessors and nonmusicians do. As a third group, classical sound engineers are examined, musicians who are trained in an absolutely precise way on hearing acoustic differences in a sound.

40 subjects from 20y to 35y (20 musicians studying a classical instrument, 5 musicians studying classical sound engineering and 15 non musicians) are tested on their auditory acuity abilities. The participants are native speakers of German and without any hearing or speaking restrictions.

For measurements of auditory acuity two continua, a sibilant and a vowel continuum, have been implemented. The sibilant continuum was the same as in Brunner, Ghosh, Hoole et al. (2011). The ends of the continuum are synthetic productions of the words “Asse” (/asə/, *ash)* and “Asche” *(*/aʃə/, *aces)*. It was created by shifting the spectral peaks stepwise, resulting in 1513 tokens. For the investigation of vowels, the continuum from Gluth and Hoole (2015) was adapted and slightly revised. The ends of this continuum were productions of the words “Beagle” (/bi:gl/, *beagle)* and “Bügel” (/by:gl/, *bracket)*, which were morphed (using the methods in Kawahara, Morise, Takahashi et al., 2008) to another set of 1513 tokens. For each continuum, two tests will be conducted: first, a labeling test, in which the subject’s phonemic boundary between the two words (*Asse-Asche* or *Beagle-Bügel*) is identified, and second, a computer-based 4-interval 2-alternative

forced choice adaptive staircase discrimination task. The implementation by Gluth & Hoole (2015) was used. For those musicians who classify themselves as absolute pitch possessors a third test was designed to ensure their absolute pitch possessing abilities by presenting randomized pure tones between a1 and a3 with white noise after each tone.

First results show differences between the groups. Detailed results will be presented at the conference.

References

[1] Brunner, J., Ghosh, S., Hoole, P., Matthies, M., Tiede, M. and Perkell, J. (2011). "The influence of auditory acuity on acoustic variability and the use of motor equivalence during adaptation to a perturbation". *Journal of Speech, Language, and Hearing Research* 54(3): 727-739.

[2] Gluth, C. and Hoole, P. (2015). *How can speech production skills be predicted from visual, auditory, and haptic perception skills?* Proc. ICPhS. Glasgow

[3] Kawahara, H., Morise, M., Takahashi, T., Nisimura, R., Irino, T. and Banno, H. (2008). Tandem-STRAIGHT: A temporally stable power spectral representation for periodic signals and applications to interference-free spectrum, F0, and aperiodicity estimation. Proceedings of the ICASSP 2008, Las Vegas 3933–3936.

[4] Kishon-Rabin, L., Amir, O., Vexler, Y. and Zaltz, Y. (2001). "Pitch discrimination: Are professional musicians better than non-musicians?". *Journal of basic and clinical physiology and pharmacology* 12(2): 125-144.

[5] Masataka, N. (2011). "Enhancement of speech-relevant auditory acuity in absolute pitch possessors". *Frontiers in psychology* 2:1-4.

[6] Oechslin, M. S., Meyer, M. and Jäncke, L. (2010). "Absolute pitch – Functional evidence of speech-relevant auditory acuity". *Cerebral Cortex* 20(2): 447-455.

[7] Pape, D. (2009). "Microprosodic Differences in a Cross-linguistic Vowel Comparison of Speech Production and Speech Perception". Weißensee-Verlag.

An automatic chunk segmentation tool for long transcribed speech recordings

Nina Poerner[1]*, Florian Schiel*[1]

[1]Bavarian Archive for Speech Signals,
Ludwig-Maximilians-Universität München
npoerner@phonetik.lmu.de, schiel@phonetik.lmu.de

Forced alignment tools such as the Munich Automatic Segmentation System (MAUS) [1] do not scale well with input size. In this paper, we present a preprocessor chunk segmentation tool to combat this problem. It dramatically decreases MAUS's runtime on recordings of duration up to three hours, while also having a slightly positive effect on segmentation accuracy. We hope that this tool will advance the use of non-scientific transcribed recordings, such as audio books or broadcasts, in phonetic research. The chunker tool will be made available as a free web service at the Bavarian Archive for Speech Signals (BAS) [2].

☞ **See full paper in Proceedings**

Geflüsterte Angst und behauchte Trauer – Stimmqualität und Emotionen

Louise Probst, Angelika Braun

Universität Trier
s2lodirk@uni-trier.de, brauna@uni-trier.de

Der Parameter Stimmqualität ist im Bereich der phonetischen Erforschung von Emotionen in der Stimme selten betrachtet worden. Die vorliegende Studie beschäftigt sich daher mit der Frage, welche Stimmqualitäten Sprecher zur Porträtierung von Emotionen einsetzen. Hierfür wurden emotionale und neutrale Sprachäußerungen von sechs professionellen Sprechern des Deutschen aufgezeichnet und hinsichtlich ihrer Stimmqualität beurteilt. Während die Vorhersagen für das Auftreten von breathy voice bei Freude zutreffen, lassen sich die Vorhersagen für harsh voice bei Wut nur teilweise und für creaky voice bei Trauer nicht bestätigen.

☞ **See full paper in Proceedings**

Open data for speech synthesis of Austrian German language varieties

Michael Pucher[1], Michaela Rausch-Supola[1], Sylvia Moosmüller[1], Markus Toman[2], Dietmar Schabus[3], Friedrich Neubarth[3]

[1]Acoustics Research Institute (ARI), Austrian Academy of Sciences (OAW)
[2]Vienna University of Technology (TUW), Austria
[3]Austrian Research Institute for Artificial Intelligence (OFAI)
{michael.pucher|michaela.rausch-supola|sylvia.moosmueller}@oeaw.ac.at,
m.toman@neuratec.com, {friedrich.neubarth|dietmar.schabus}@ofai.at

In this paper we summarize open data sets and open source software that we have released for Austrian German language varieties as a result of several research projects. We also describe some data sets that are released for research purposes only, due to licensing limitations. From the development of these resources we draw conclusions concerning the collection and licensing of such data with a special focus on the problem of speech synthesis where the voice identity of the speaker plays an important role. Furthermore we discuss recordings that we plan to perform in the future, where we aim to cover most Austrian dialects.

☞ **See full paper in Proceedings**

Romance Corpus Phonology: from (Inter-)Phonologie du Français Contemporain (I)PFC to (Inter-)Fonología del Español Contemporáneo (I)FEC

Elissa Pustka[1], *Christoph Gabriel*[2], *Trudel Meisenburg*[3]

[1]Institut für Romanistik, Universität Wien,
[2]Universität Mainz, Germany
[3]Institut für Romanistik/Latinistik, Universität Osnabrück, Germany
elissa.pustka@univie.ac.at, christoph.gabriel@uni-mainz.de, tmeisenb@uos.de

Contemporáneo (I)FEC aims to document the pronunciation of Spanish in the world, including L1 and L2 speakers as well as learners of Spanish as a foreign language. Our starting point is the French research program *(Inter-)Phonologie du Français Contemporain* (I)PFC. On the basis of its nearly 20 years of experience and several pilot studies on Spanish, we present for the first time the guidelines we developed for the data collection in this project. (I)FEC works with a modular system: the basic design elicits data via a reading task (word list and text), a discourse completion task and a semi-focused interview. For some types of speakers (learners, illiterates, multilinguals), we provide supplementary tasks. In doing so, we take into account variation in both segmental and suprasegmental phenomena such as regionally confined oppositions (e.g., /s/:/θ/), the weakening of coda consonants (particularly coda /s/), word stress, syllabification and intonation.

☞ **See full paper in Proceedings**

AI vs. AU in American English compared to German

Renate Raffelsiefen, Anja Geumann

Institut für Deutsche Sprache, Mannheim
raffelsiefen@ids-mannheim.de, geumann@ids-mannheim.de

American English and German AI, AU observed in cognates such as *Wein, wine, Haus, house* are usually treated on a par, represented with the same initial vowel (cf. [aɪ], [aʊ] for Am. Engl. and German [1]). Yet, acoustic measurements indicate differences as the relevant trajectories characteristically cross in Am. Engl. but not in German. These data may indicate consistency with the same initial target for these diphthongs in German, supporting the choice of the same symbol /a/ in phonemic representation, as opposed to distinct targets (and distinct initial phonemes) in American English.

☞ **See full paper in Proceedings**

Acoustic and articulatory manifestations of final lengthening and voicing contrasts for German learners of English as a second language

Oxana Rasskazova [1], Malte Belz [1] , Christine Mooshammer [1,2], Jelena Krivokapić [2,3]

[1] Humboldt-Universität zu Berlin, Germany
[2] Haskins Laboratories, USA
[3] University of Michigan, USA
oxana.rasskazova@hu-berlin.de, malte.belz@hu-berlin.de

German native speakers show prosodic and segmental transfer effects of the L1 when speaking English as L2. The focus lies on prosodic lengthening patterns as well as voicing contrast in word-final position in English, a pattern that is difficult for many German learners due to their native syllable-final obstruent devoicing rule.

☞ **See full paper in Proceedings**

What is the fate of Scottish Vowel Length Rule in Glasgow?

Tamara Rathcke[1], *Florent Chevalier*[2], *Jane Stuart-Smith*[3]

[1] English Language and Linguistics, University of Kent
[2] Modern Languages and Cultures, University of Glasgow
[3]English Language, University of Glasgow
t.v.rathcke@kent.ac.uk, florent.chevalier@glasgow.ac.uk, jane.stuart-smith@glasgow.ac.uk

This paper studies the longitudinal development of a vowel timing alternation known as the "Scottish Vowel Length Rule" in a distinctive variety of Scottish English spoken in Glasgow by working-class men and women. Combining apparent-time and real-time evidence, we show that the implementation of the Rule has changed over time, though unlike in many other varieties of Scottish English, the factors shaping its fate seem to be internal rather than external. Overall, Glaswegian English behaves like a quantity language and controls for prosodic timing effects while preserving the phonological timing alternation; and this is despite a marginal, quasi-phonemic status of the Rule.

☞ **See full paper in Proceedings**

Entrainment analysis of categorical intonation representations

Uwe D. Reichel[1], *Jennifer Cole*[2]

[1]Research Institute for Linguistics, Hungarian Academy of Sciences, Hungary
[2]Department of Linguistics, Northwestern University, United States of America
uwe.reichel@nytud.mta.hu, jennifer.cole1@northwestern.edu

Most studies on prosodic entrainment focus on coarse parametric variables as f0 mean and standard deviation. Only recently first attempts were made to measure entrainment also for categorical intonation representations namely pitch accent types [1]. We propose further metrics for this purpose adopted from text similarity measurement and alignment. These metrics were applied to quantify the similarity of automatically derived intonation contour class sequences in cooperative and competitive dialogs. In line with previously reported results for parametric variables we found also for the categorical representation higher similarities and thus more entrainment in the cooperative dialogs than in the competitive ones. The introduced metrics can be of use for any entrainment research on categorical data as e.g. for ToBI label sequences.

☞ **See full paper in Proceedings**

An acoustic analysis of German initial laterals in the L2 speech of Bosnian migrants living in Vienna

Carolin Schmid

Acoustics Research Institute, Austrian Academy of Sciences, Vienna
carolin.schmid@oeaw.ac.at

In this paper, the phonetic realization of L2 German initial laterals produced by L1 Bosnian speakers living in Vienna is investigated and compared to L1 Standard Austrian German (SAG) and L1 Viennese Dialect (VD) realizations, both in read speech and in spontaneous speech. This pilot study is embedded in the larger context of my PhD dissertation on the sociophonetic aspects of language contact in the speech of Bosnian migrants living in Vienna, which will be concerned with language acquisition as well as language attrition. The results show that Bosnian speakers realize a more velarized lateral in initial word position than both SAG and VD speakers, both in read and in spontaneous speech and in all vowel contexts. Thereby, the velarization is stronger in read speech than in spontaneous speech.

☞ **See full paper in Proceedings**

Wie Deutschschweizer Lernende die stimmhaften Obstruenten des Italienischen aussprechen

Stephan Schmid

Phonetisches Laboratorium, Universität Zürich (Schweiz)
stephan.schmid@uzh.ch

Im Gegensatz zum Italienischen gibt es in Deutschschweizer Dialekten keine stimmhaften Obstruenten, sondern (stimmlose) *lenes*, die sich von den (ebenfalls stimmlosen) *fortes* hauptsächlich in der Dauer unterscheiden. Aufgrund der gängigen Modelle der L2-Phonologie ist anzunehmen, dass deutschschweizer Lernende stimmhafte Obstruenten des Italienischen an die *lenes* ihrer L1 'assimilieren'; eine kontrastive Analyse legt zudem nahe, dass dies bei /dz/, /dʒ/ und s+CVerbindungen besonders häufig der Fall sein könnte.
Diese Hypothesen wurden im Wesentlichen durch eine akustische Analyse von 644 italienische Obstruenten bestätigt, die von 10 Zürcher Mittelschülern realisiert wurden. Dabei ergaben sich gewisse Unterschiede zwischen den einzelnen Sprechern, aber auch aufgrund der Faktoren 'Konsonant' und 'Kontext'; letztere können z.T. anhand von Markiertheitsüberlegungen erklärt werden.

☞ **See full paper in Proceedings**

Untersuchung des Kompensationsverhaltens bei Echtzeitmanipulation der Zeitstruktur des auditorischen Feedbacks

Laura Sichlinger, Philip Hoole

Institut für Phonetik, LMU München
laura.sichlinger@campus.lmu.de

Auditorisches Feedback ist ein elementarer Bestandteil der Sprachproduktion. Studien mit Echtzeitmanipulation zeigten, dass die Integration der auditorischen Rückmeldung in ein Kontrollsystem für Sprechbewegungen ausschlaggebend für deren Planung ist. Derartige Manipulation resultierte in artikulatorischen Kompensationsmechanismen. Frühere Untersuchungen beschränkten sich allerdings auf spatiale Manipulationen des Sprachsignals. Da die zeitliche Organisation des Sprechens aber eine beträchtliche Auswirkung auf die Produktion von Sprache hat, ist es unumgänglich diese Komponente in Verbindung mit der auditorischen Rückmeldung und Integration von dieser zu untersuchen. Daher wurde ein Experiment durchgeführt, in dem komplexe Stimuli zeitlich perturbiert und die artikulatorische Reaktion der Sprecher analysiert wurde. Es wurde zudem überlegt, dass aufgrund von engeren zeitlichen Beziehungen innerhalb des Onsets einer Silbe im Vergleich zur Coda eindeutigere Anpassungsmechanismen in der artikulatorischen Reaktion des Sprechers bei Onsetmanipulierten Sprachsignalen entstehen. Dafür wurden zwei Stimuli erstellt, von denen jeder jeweils einer Gruppe präsentiert wurde. Sowohl Onset- als auch Coda-manipulierte akustische Signale zogen artikulatorische Kompensationsmechanismen nach sich. Außerdem wurden Sprechbewegungen stärker in Reaktion auf Onset-Manipulationen verändert, was darauf hindeutet, dass zeitliche Kontrollmechanismen in der Sprachproduktion der Perzeption unterliegen und dass Lautstrukturen innerhalb des Onsets im Vergleich zur Coda abhängiger von der mentalen Vorhersage der Zeitstruktur eines Lauts sind.

☞ **See full paper in Proceedings**

Prosodic variation in conceptual distance and proximity: Self-repairs in French

Johanna Stahnke

Bergische Universität Wuppertal, Germany
stahnke@uni-wuppertal.de

The present study addresses the prosodic contextualization of self-repair in conceptual distance and conceptual proximity in spoken French. According to the literature, paraphrases establish semantic equivalence and are intonationally deaccented; corrections constitute semantic difference and show overaccentuation. Repairs are expected to be more frequent in spontaneous proximity than in pre-planned distance. Apart from quantities, it is hypothesized that selfrepair is subject to qualitative differences: As a consequence of contextual involvement in conceptual proximity, speakers should more extensively use prosodic contextualization cues than in context-detached distance. Rather than being contextualized for repair type, however, the results of an empirical analysis of two conceptually contrastive corpora suggest a general tendency of deaccentuation in proximity and of overaccentuation in distance, independently of the repair type. The interpretation of these unexpected findings is proposed in a communicative model of language variation and change, in which speakers in conceptual proximity strategically routinize deaccented structures in order to cope with conversationally undesired disruptive corrections. Possibilities for prosodic change are discussed in the light of lexicalization of discourse markers induced by conceptual proximity.

☞ **See full paper in Proceedings**

Vowel confusions in noise by German listeners: A study of oral and nasalized vowels

Kim Strütjen, Ruben van de Vijver

Institut für Sprache und Information, Heinrich-Heine-Universität Düsseldorf, Deutschland
Kim.Struetjen@hhu.de, Ruben.Vijver@hhu.de

Low oral and nasalized vowels are acoustically more similar than non-low oral and nasalized vowels [1], [2]. We hypothesize that low oral and nasalized vowels are more likely to be confused with each other than non-low oral and nasalized vowels, even in languages without contrastive nasalization or allophonic nasalization, such as German [3, pp. 12-13, 32].
In a forced-choice identification task 28 native German adults listened to vowels in the presence of a masking noise and were forced to identify the vowels as one of these six vowels: [a], [ã], [ɛ], [ɛ̃], [i] or [ĩ]. The results confirm the hypothesis and show that native Germans confuse [a] more often with [ã] than [ɛ] with [ɛ̃] or [i] with [ĩ]. These results are in line with previous studies indicating the tendency that high nasalized vowels are easily distinguished from high oral vowels by native speakers of American English [4], [5]. However, native American adults are familiar with allophonically nasalized vowels [6] whereas our participants are not. Our study, therefore, shows that the acoustic modification due to the nasalization is perceived better in nonlow vowels than in low vowels even by speakers who are not familiar with vowel nasalization.

☞ **See full paper in Proceedings**

Then, what is charisma? The role of audio-visual prosody in L1 and L2 political speeches

Hiroyuki Tanaka[1], *Tamara Rathcke*[2]

[1]Kotobaax, The Hague, Netherlands
[2] English Language and Linguistics, University of Kent, UK
kotobaax.hiro@gmail.com, t.v.rathcke@kent.ac.uk

Charisma plays a significant role in political speeches, and determines the ability of a politician to carry an audience. While acoustic features of charisma have received some empirical attention, the contribution of visual prosody has been mostly neglected in studies focusing on features of a charismatic appearance. Unknown are also the audio-visual cues to charisma in non-native speakers. This small-scale study investigated speeches delivered by Donald Trump (L1 American English) and Arnold Schwarzenegger (L1 Austrian German, L2 American English). Video and audio recordings of their political speeches (around 25 min per speaker) and the transcripts were used. The use of pitch range, speech rate, emphatic stress and hand gestures was analysed. In order to establish the core means of the speakers' persuasive influence on their audiences, within-speaker comparisons were conducted for phrases with and without cheering from the audiences. The results showed some differences in the use of the audio-visual prosodic features between the L1 and L2 speaker as well as some similarities, and suggest that charisma is not easily attributable to a fixed set of prosodic means but may be best understood as a skillful modulation of audio-visual prosody in social interaction.

☞ **See full paper in Proceedings**

Coordination Deficits in Essential Tremor Patients with Deep Brain Stimulation

Tabea Thies, Anne Hermes & Doris Mücke

IfL Phonetik, University of Cologne, Germany
{tabeathies|annehermes|doris.muecke}@uni-koeln.de

It has been shown elsewhere that speech is deteriorated in Essential Tremor patients treated with Deep Brain Stimulation [2]. However, those studies were restricted to fast syllable repetition tasks. The present study is a pilot study investigating the coordination of oral gestures in normal sentence production using Electromagnetic Articulography (EMA). It focuses on intergestural coordination patterns of the labial and lingual system. We test the interplay of these patterns in syllables with different complexity within the framework of Articulatory Phonology. When comparing simple (CV) and complex syllables (CCV), it is assumed that the rightmost C of a consonant cluster shifts towards the vowel whereas the leftmost C shifts away from it to make room for the added C. However, we did not find this expected shift of the rightward C in the patients' production, indicating an articulatory mistiming within a prosodic constituent such as the syllable.

☞ **See full paper in Proceedings**

Die Wahrnehmung reduzierter Sprache unter Rauschen

Frederike Urke[1], *Prof. Dr. Henning Reetz*[2], *Dr. Gea De Jong-Lendle*[1]

[1] Philipps-Universität Marburg, Deutschland
[2] Goethe-Universität Frankfurt, Deutschland
frederike.urke@staff.uni-marburg.de, reetz@em.uni-frankfurt.de, gea.dejong@staff.uni-marburg.de

Diese Studie befasst sich mit der Frage, ob es einen Unter-schied beim Verständnis reduzierter gegenüber kanonischer Sprache im Störgeräusch gibt. Das durchgeführte Experiment zeigt, dass reduzierte Sprache unter rosa Rauschen schlechter erkannt wird als die kanonischen Pendants. Grund dafür könnte das Wechselspiel zwischen der Maskierung akustischer Eigenschaften, die für das Verständnis reduzierter Sprache relevant sind, und den Reduzierungen selbst sein.

☞ **See full paper in Proceedings**

Beat it! – Gesture-based Prominence Annotation as aWindow to Individual Prosody Processing Strategies

Petra Wagner[1,2], Aleksandra Ćwiek[1], Barbara Samlowski[3]

[1]Fakultät für Linguistik und Literaturwissenschaft, Universität Bielefeld
[2]Cluster of Excellence Cognitive Interaction Technology (CITEC)
[3] Amazon Development Center, Aachen
petra.wagner@uni-bielefeld.de, aleksandra.cwiek@uni-bielefeld.de

In recent work [1], we have suggested a novel approach for fine-grained and fast prominence annotation by na¨ıve listeners. Our approach relies on annotators' "drummed" replications of a perceived utterance, modulating their drumming velocity in accordance with the perceptual prominence of consecutive linguistic units (syllables, words). The drumming velocity is then used as a fine-grained operationalization of prosodic prominence. Due to its speed and ease, it allows for the rapid annotation of large amounts of data and yields results that are comparable to fine-grained expert annotations of prominence.
In the present study, we evaluated our method further by (1) comparing the intra-sentential prosodic variation as measured with traditional annotations and the drumming method. Our results show that "drummed" prominences capture speaking-stylerelated variability similarly to conventional annotation methods. Additionally (2), we examined whether individual processing strategies can be identified with the help of Random Forests. This method allows for estimating the individual impact of established prominence correlates on prominence impressions. Our analyses unveil individual listener strategies for blending and integrating top-down, bottom-up and context cues into impressions of prosodic prominence.

☞ **See full paper in Proceedings**

Two-stage Decision Trees for Automatic Speaker Likability Classification

Mathias Walther[1], Oliver Jokisch[2], Taïeb Mellouli[1]

[1] Martin Luther University Halle-Wittenberg, Germany
[2] Leipzig University of Telecommunications, Germany
mathias.walther@wiwi.uni-halle.de

This article discusses a two-stage classification system for paralinguistic speaker traits which is part of a prototypical expert system for rhetorical and vocal quality assessment in call center talks. The system is based on pre-trained models for vocal features and outputs comprehensible classification rules so that the agent can improve his rhetorical abilities. The recognition of vocal features is modeled by competing classification systems and combined into a multi-classifier system which is based on decision trees. We compare two decision tree inducers, namely C4.5 and random forest, both in prediction accuracy and their rule sets. The experiments were conducted with the Speaker Likability Database (SLD) and benchmarked against the results of the Interspeech 2012 Speaker Trait Challenge. In terms of accuracy, the proposed two-stage classification performs similar to the baseline results with the advantage of being introspectable.

☞ **See full paper in Proceedings**

Changes in IDS and ADS during parental leave – project sketch and first results of pilot study

Melanie Weirich, Adrian Simpson

Institute of German Linguistics, Friedrich Schiller University Jena
Melanie.weirich@uni-jena.de, adrian.simpson@uni-jena.de

We are investigating the speech of German and Swedish mothers and fathers during the first year of their first baby. Both infant- and adult-directed speech are analyzed and compared between the sexes but also between different time points during the first year. In addition, the involvement in child care is considered as a potential factor. We are now in the process of finding participants and gathering data (read speech and spontaneous speech) from the first recording before birth of the child. Here, the speech material, our hypotheses and first pilot studies are presented.

☞ **See full paper in Proceedings**

Voice Descriptions by Non-Experts: Validation of a Questionnaire

Benjamin Weiss

Technische Universität Berlin, Germany
benjamin.weiss@tu-berlin.de

A questionnaire for non-expert German listeners was developed and validated in order to describe unacquainted voices on common perceptual dimensions. This instrument can be used to study acoustic correlates of perceptual features and voice-based attributions. Thirteen male and thirteen female speakers have been rated on a comprehensive questionnaire with 34 bipolar items by samesex listeners (30 women, 31 men). Two stimuli of each speaker have been evaluated. The aim was to finalize the development of the questionnaire by internal validation with factor analysis, and by testing for correlations with meaningful acoustic parameters for each obtained factor. After disregarding only few questionnaire items, the final sets of items show very good performances and reveals five, respectively six factors: Activity, Fluency, Precision, Softness, and Darkness for both, and additionally Tempo for women. Acoustic parameters were extracted automatically from the speech signal. Applying step-wise inclusion of preselected parameter sets, significant linear models were obtained for most of the factors. Only pronunciation Precision and Fluency will require more sophisticated approaches for automatic analysis.

☞ **See full paper in Proceedings**

The impact of accent familiarity on the perception of difficult sound contrasts for German learners of English

Katrin Wolfswinkler[1], *Eva Reinisch*[1]

[1]Institute of Phonetics and Speech Processing, LMU Munich
k.wolfswinkler@yahoo.de, evarei@phonetik.uni-muenchen.de

Second language (L2) learners usually retain a foreign accent even after years of training. The present study aimed at finding reasons for why accents are so persistent. We hypothesized that familiarity with an accent due to frequent exposure leads to adaptation which in turn allows listeners to understand the accent better, reducing the need for improvement. If this was the case, L2 learners should be better at understanding words spoken in a familiar than unfamiliar accent. To test this account, English minimal pairs containing two difficult sound contrasts for German learners (/E/ vs. /æ/; voiceless vs. voiced wordfinal stops) were presented to native German listeners who had to identify the intended word. The tokens were produced by native speakers of English who served as a control, fellow Germans whose accent was highly familiar, Italians whose accent was somewhat familiar, and Finnish whose accent was unfamiliar. Results showed that words spoken by native English speakers or fellow Germans were recognized best, but words with the unfamiliar Finnish accent were recognized better than Italian accented words. A closer analysis of the acoustic differences that our speakers produced suggests that the acoustic cues available in the speech signal determined how well the words were identified.

☞ **See full paper in Proceedings**

'Chend' met <e> – 'Kind' mit <e>: using Big Data to explore phoneme-to-grapheme mapping in Lucerne Swiss German

Urban Zihlmann, Adrian Leemann

University of Cambridge
{ubz20|al764}@cam.ac.uk

Speakers from the canton of Lucerne are infamous for spelling Middle High German (MHG) <i> as <e> when communicating in written Swiss German, e.g. *Kind* ('child') as <Chend>. This phenomenon has been examined only impressionistically by phoneticians. This study provides a first account of this peculiarity of Lucerne Swiss German spellers: an analysis of normalised formant frequencies of two underlyingly MHG <i> vowels from 200+ speakers of the *Dialäkt Äpp* corpus revealed that the Lucerne allophone is in reality [e] for most of the localities examined, which may explain why in vernacular writing, spellers prefer <e> over <i>. Homophony due to this peculiarity can cause misunderstandings in written and oral communication, and possibly has repercussions on the reading and writing development of Lucerne students.

☞ **See full paper in Proceedings**

STICHWORTVERZEICHNIS