

**Part 4: HARMONY – SEGMENTS
IN CONTEXT**

**Część 4: HARMONIA – SEGMENTY
W KONTEKŚCIE**

NEUTRALIZATION.?! The phonetics-phonology issue in the analysis of word-final obstruent voicing

NEUTRALIZACJA.?! Zagadnienie fonetyki-fonologii w analizie dźwięczności wygłosowych obstruentów

Klaus J. Kohler

University of Kiel, Germany
kjk@ipds.uni-kiel.de

ABSTRACT

In German and Polish, the voicing contrast in obstruents is excluded from word-final position. Prague Circle phonology treated this as phonological neutralization with phonetic voiceless exponency. Generative Phonology created underlying phonological contrasts, which experimental phoneticians started testing for phonetic substantiation at the surface level. Laboratory Phonology has adopted this paradigm and continues to maintain that neutralization is incomplete in spite of differences being too small for communicative relevance. The debate can be terminated by substituting a polysystemic syntagmatic approach for the monosystemic paradigmatic one.

STRESZCZENIE

W języku niemieckim i polskim w obstruentach w pozycji wygłosowej w wyrazie nie istnieje kontrast pod względem dźwięczności. Fonologia w ujęciu Koła Praskiego traktowała to jako neutralizację fonologiczną z bezdźwięczną manifestacją fonetyczną. Fonologia Generatywna stworzyła leżące u podstaw zjawisk kontrasty fonologiczne, które fonetycy eksperymentalni zaczęli testować pod kątem fonetycznych korelatów na poziomie powierzchniowym. Fonologia Laboratoryjna przyjęła ten paradygmat i nadal utrzymuje się, że neutralizacja jest niekompletna, mimo że różnice są zbyt małe, aby miały znaczenie dla komunikacji. Kres debacie może położyć zastąpienie podejścia wielosystemowego i syntagmatycznego podejściem jednosystemowym i paradygmatycznym.

1. From the Archiphoneme to Incomplete Neutralization

In a fair number of languages, the voicing distinction in obstruents is limited to certain positions. In Polish, the domain of these limitations is the phrase, across morpheme and word boundaries: all final obstruents are voiceless before pause, and internal obstruent clusters are either all voiceless or all voiced depending on the voicing feature of the last obstruent [1]. In German, the domain is the word: all obstruents are voiceless in final position (also morpheme-final before consonant, e.g. *möglich* /k/ vs. *hüg(e)lig* /g/) [2]. These phonotactic restrictions were conceptualized in Prague Circle phonology as neutralization of a phonological opposition to an archi-phoneme, which incorporates all the features that the elements of the neutralized pair have in common, and thus does not coincide with either of them, but which is represented phonetically

by one or the other [3]. In the cases quoted, it is the voiceless obstruent in German, the voiceless or the voiced one in Polish.

For alternations between archiphoneme and voiced or voiceless obstruent in morphological paradigms, as between German *Bund* [t] – *Bunde*, *Bünde* [d], *bunt* [t] – *bunte* [t], Prague Circle linguistics, just like American Structuralism, set up a morpho-phonological level, separate from phonology in the narrow sense. This conception of levels of linguistic description changed with the advent of Generative Grammar, which denied systemic phonology an independent status and integrated phonological correspondences into morphological paradigms in the phonological component of a grammar. Phonological alternations were represented by the same contrast of underlying distinctive feature bundles across all positions, from which the neutralized form was derived by phonetic rule, e.g. word-final devoicing of /d/ in *Bund* [t]. Thus neutralization was treated as a phonetic process operating on a phonological distinction.

So, when researchers started taking these new phonological solutions into the laboratory, it was a natural consequence for them to ask whether the postulated non-neutralization in *phonology* was also supported by a *phonetic* contrast, which had not been detected previously due to missing measurements. The first to raise this incomplete neutralization issue for German were Mitleb [4], Port et al. [5], O'Dell and Port [6], followed by more extensive accounts and discussions in Port and O'Dell [7]. These researchers stood in the Haskins Labs tradition of analysing the voiced/voiceless contrast in English word-final plosives with reference to the durations of the preconsontal vowel, the closure and closure voicing, and the release, which they applied to German words. Consistent differences were found, especially in the duration of the vowel, in citation form pronunciation. They were small (15 ms of vowel and burst durations, 5 ms of closure voicing on average), but in the same direction as in the non-neutralizing word-medial context, i.e., longer vowel, shorter burst, and more closure voicing in the underlying voiced plosive, some of them reaching statistical significance. The production data were supplemented by listener judgments, which identified the produced member of a pair poorly, but significantly better than chance (59% overall). The authors concluded from these results, supported by inferential statistics, that German had incomplete neutralization of the underlying word-final voicing opposition in obstruents, and that these phonetic differences would have to be taken into account in the phonology of German by the introduction of implementation rules of the generally maintained phonological contrast.

Fourakis and Iverson [8] argued convincingly, on the basis of data from orthographic vs. free elicitation of relevant words, that the phonetic differences were caused by hyper-correction of isolated words in a reading task, and that there were no grounds for postulating incomplete neutralization in the phonology of German. They also pointed out that *weg* (although etymologically related to the noun paradigm *Weg* : *Wege*) was not part of a morphological alternation at all, now being an adverb with a short vowel. Another crucial point in the evaluation of these data is that the authors found it necessary to explain the meanings of some of their words to their subjects before the actual tests, viz. *Alb* 'elf' vs. *Alp* 'mountain pasture', *schräg* (imperative form of obsolete *schrägen*) 'join wooden poles cross-wise or slanting' vs. *schrak* (past tense of intransitive *schrecken*, which only occurs in *erschrecken*) 'had a fright', *Weck* 'breakfast roll' (southern dialect word). In view of the complete lack of context for words, which in addition are not all

part of the subjects' common vocabulary or are unusual morphological forms, the test items assumed the status of logotomes. Furthermore, the data were collected in the US from German speakers who had spent considerable lengths of time in an English-speaking environment and had no doubt been made aware of the need to differentiate final orthographic , <d>, <g> from <p>, <t>, <k>. In view of these flaws in experimental design, these data were not able to substantiate the claim of incomplete neutralization. But the discussion has continued ever since.

Slowiaczek and Dinnsen [9] applied the same phonological paradigm with the same experimental design flaws to neutralization in Polish, and obtained the same type of small differences. On the basis of such phonetic measurements in underlying phonological distinctions, Dinnsen [10, 11] finally concluded that careful phonetic studies would reveal that all declared neutralization rules in the phonologies of languages were in effect not neutralizing at all. Not only did he accept the validity of the phonetic data without query, he also put forward theoretical phonological arguments to counteract the objection that the acoustic production differences are too small to be perceived. As in generative phonology grammars are assumed to be abstract representations, neutral with respect to speaker and hearer, the existence of distinctions in production but of neutralization in perception would entail a different theoretical conceptualization. The speaker-hearer indifference would vanish, neutralization would become part of phonological perception, and production differences would be linguistically irrelevant. This would have "awkward consequences" for allophonic rules, which "describe production differences [that] are not generally self-discriminable by native listeners" So, if "phonological rules specify segments differently only if their production differences are perceptually salient, phonological descriptions would exclude all allophonic statements" [11:pp271f].

There are two flaws in this line of argument. First of all, it mixes up non-discriminability on the basis of an acoustic distinction being below the Just Noticeable Difference (JND) of the perception mechanism, and non-discriminability of contextual variation in category allocation experiments. Naïve listeners in most English varieties may not be able to 'hear' the difference they make between clear and dark /l/ initially and finally in English *little*, when asked about the presence of a sound distinction, but they will immediately recognise the absence of the right /l/ in final position and categorize it as 'foreign' or 'Irish'. The lack of perceptual relevance of acoustic measurements in the incomplete neutralization issue, however, has to do with the JND threshold. The second flaw concerns the conception of a language as an abstract system indifferent as to speaker and hearer, and thus dissociated from speech communication, but nevertheless buttressed with reference to speaker data obtained in problematic experimental lab speech designs.

2. Refuting Incomplete Neutralization in Polish

Jassem and Richter [1] replied to [9] with an experimental analysis that removed the flaws of previous data acquisition and placed phonological description in a framework of speech communication rather than on a level of abstract linguistics. The findings of [9] were reexamined with materials that were to meet the following conditions:

- speakers are recruited within their native-language country,
- speakers do not see the materials in orthographic form,
- materials are not of a metalinguistic type but consist of unprepared utterances elicited in a natural dialogue situation,
- speakers represent a homogeneous group with regard to dialectal variation.

The results show a slight tendency of mean durations of vowels towards greater, and of consonants and consonant releases towards smaller values for +voice. The differences are extremely small and not always regular (4 ms for vowel and for consonant durations), and none of the differences reach statistical significance. In a subsequent listening test, subjects indicated a strong bias for phonetic voicelessness corresponding to underlying –voice as well as +voice in final neutralizing position. So the conclusion is that “data obtained from materials recorded under conditions approaching a spontaneous discourse situation give no ground for rejecting the traditional view that in positions of phrase-final devoicing and internal assimilation, neutralization of the voice feature in Polish obstruents is phonetically (and *eo ipso* phonologically) complete. ... Extrapolating from experiments performed earlier elsewhere under conditions of lesser naturalness than ours, as well as from our own results, there appears good reason to conclude, as a generalization, that if any linguistic-phonetic (phonological) descriptions are to be tested experimentally, it is mandatory to ensure *maximum naturalness* of the materials.” [1:p124]

3. The continued issue of Incomplete Neutralization in German

In spite of the critical assessment of the research into incomplete neutralization of voice in final obstruents (see also [12] for a critique of the German data), researchers persisted in pursuing the question, with little or no regard to Wiktor Jassem’s mandatory guidelines for data acquisition. Latterly the Laboratory Phonology framework became the ideal environment to carry phonological solutions of final voicing restrictions into the laboratory for phonetic substantiation and thus kindle renewed investigation into incomplete neutralization. A recent paper [13] is part of this paradigm.

In a production experiment, subjects successively heard 48 randomized non-words (disyllabic minimal pairs with intervocalic /bp/ (8), /dt/ (7), /gk/ (9) of German phonotactic structure), each embedded in a meaningful German sentence, pronounced by a male speaker, in which the non-word was decodable as a masculine plural noun. They were asked to respond to each heard stimulus sentence by a given sentence frame in which they were to place the noun in the singular, thus moving the plosives into final position. 16 native speakers of German from the Cologne area, 9 women, 7 men, participated in the experiment. In the disyllabic prompts, the vowel was on average 28 ms shorter before voiceless stops, closure duration and VOT were on average 21 ms and 41 ms longer, all statistically highly significant. So, the acoustic differences in the disyllabic pairs were strongly marked. In the monosyllabic responses, however, only vowel duration reached significance, and on average was only 8 ms shorter before the voiceless plosive. This minimal value is below the JND, and there are thus worlds between the signalling of the underlying phonological contrasts in the two contexts.

The authors maintain that subjects could not be influenced by orthography in the production experiment since the task did not include written words but only aural stimuli which had to be output in modified form. However, since the subjects were no doubt undergraduates as well as PhDs from the Cologne Linguistics Department, a fairly high degree of literacy can be assumed, even to the extent of a conscious awareness of the relationship of sounds and letters. Furthermore, when naive language users are to repeat real spoken words they recognise them as such and are then able to output them again. If they also have to modify them this already introduces an interfering variable. But if they cannot categorize the auditory stimuli linguistically as existing words they will find it very difficult to even just repeat them. Categorization via an orthographic representation will help them, but this instrument is only available to literate people and works the more efficiently the greater the degree of literacy and awareness of the sound – letter relationship. The fact that a considerable 12% of the data (16 * 48 = 768 items), i.e. an average of 5-6 items per speaker, had to be excluded because they were either incorrectly remembered or reproduced with a lot of hesitation, points in the same direction. So, orthographic influence cannot be confidently excluded.

In the subsequent perception experiment, 8 listeners had to judge 192 sentence items from the production experiment by referring to the orthographic alternatives presented for each word (e.g. <Drud> vs. <Drut>). They were instructed that exactly half of the stimuli were from the set <b, d, g> and half were from the set <p, t, k>. This introduced a bias towards voicing in the experimental design despite the absence of strong cues for it in the production data, and also raised the number of false alarms in voiceless items. The overall average accuracy of voiced or voiceless identification was 54%, pointing to high proportions of misses and false alarms. Participants ranged from 50% to 59%. This is rather poor performance, compared with 98% correct identification of the prompts. In view of these results, subjects' judgments cannot be based on voiced/voiceless cue perception but are random responses, in spite of statistical significance of the small vowel duration difference. This means that the existence of incomplete neutralization in German has not been proved in this renewed perceptual investigation either.

It is thus unjustified to conclude "...our results demonstrate that IN effects in perception are robust to the extent that they occur in a forced-choice paradigm even with a multitude of voices ... These experiments therefore address important concerns with previous investigations of incomplete neutralization. Our results indicate that IN is a robust phenomenon whose phonological implications should be taken seriously." [13:p1725]

4. A different approach to Neutralization

In this Laboratory Phonology approach, phonetic data are no longer analysed and interpreted as sound patterns in speech communication but are related to abstract phonological categories which are removed from speech and have their aim in themselves, and inferential statistical tests are applied to validate them. This is what the investigation into incomplete neutralizations is all about. And since their acoustic manifestations and perceptual consequences are so weak and unclear, and based on

flawed experimental designs far removed from natural speech communication, they cannot be postulated as having any relevance in human interaction by sound.

It is high time that researchers in speech science heard Wiktor Jassem's lone cry from the phonetic wilderness, and stopped wasting enormous resources of time, manpower and public money chasing a phonology-induced phantom, and that they turned their scientific skills instead to the elucidation of more interesting questions of real speech communication. A look at the problems German children have learning to write can convince anybody, without sophisticated speech signal analysis, that, e.g., *Kind* sounds like *Kint* to them, because that is how they spell it before they learn the morphological rules.

Neutralization of the voicing feature in word-final obstruents in Standard German and its associated dialects is a phonetic reality in speech interaction, just as it is in Polish and no doubt other languages, whereas incomplete neutralization is an experimental artifact in the wake of generative phonology. But in colloquial German the phonetic reality goes even beyond the word and is becoming a feature of the phrase, operating across word boundaries, as was described *mutatis mutandis* for Polish by Wiktor Jassem. In formally and semantically cohesive lexicalized phrases, such as *außer Rand und Band*, *auf Schritt und Tritt*, *Lug und Trug*, *Dick und Doof*, *lieb und teuer*, *ab und zu*, the realization of the final plosive in the first word fluctuates along a VOT scale from voiceless aspirated to unaspirated to voicing, irrespective of the phonological voicing feature of the plosive in the morphological paradigm, only depending on the syntagmatic cohesion of the utterance in spontaneous communication. If the speaker breaks the articulatory flow after the first word by sentence-prosodic means – a tonal boundary, articulatory deceleration, a pause, glottal onset of the following vowel – the realization is voiceless aspirated. The greater the syntagmatic prosodic cohesion between the words in the phrase, the greater the likelihood that the plosives lose their aspiration and become phonetically voiced, no matter whether they are “phonologically” /bdg/, as in *Rand* vs. *Ränder*, *Zug* vs. *Züge*, and *lieb* vs. *liebe*, or /ptk/, as in *Schritt* vs. *Schritte*, *dick* vs. *dicke*, and *ab*, which is outside a paradigm, and whose final plosive must today be allocated to /p/, in spite of OHG *aba*, MHG *ab(e)*.

Figure 1 gives some examples produced with varying degrees of cohesion in a pilot production test, where subjects had to read a list of sentences that contained such lexicalized phrases, in which the first word was replaced by ____, to detract from orthographic representation of the target word, e.g. *Er gerät nicht selten außer ____ und Band. Er folgt ihr auf ____ und Tritt. Wir erledigen das Geschäft ____ um Zug. Laurel und Hardy sind bei uns als ____ und Doof bekannt. Sein Job ist ihm ____ und teuer. Ich schlage vor, dass wir uns ____ und zu in der Mensa treffen.* Subjects had to repeat each sentence several times.

These syntagmatic processes spanning the exponents of a paradigmatic voicing contrast can be related to a rhythmical principle that is operative in such fixed dual-element phrases [14]. In Germanic languages, their rhythmicality is governed by the number of syllables (in English *Stan and Ollie* is more rhythmical than the reverse), or by high front vs. back/low, and short vs. long vowel (the German equivalent *Dick und Doof* is more rhythmical than the reverse, and reverses the order of physical appearance in *Laurel and Hardy*). Friction noise in the first part of such phrases interrupts the

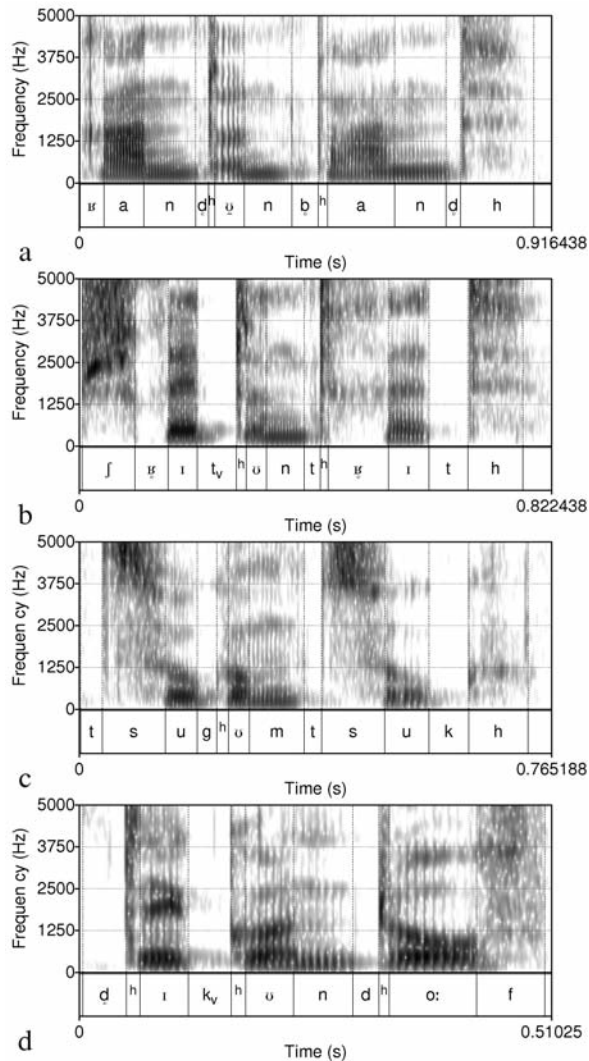


Figure 1: Spectrogram and phonetic segmentation of the German phrases
a “Rand und Band”, b “Schritt und Tritt” (female speaker),
and c “Zug um Zug”, d “Dick und Doof” (male speaker).

rhythmic flow. That is an additional reason why *Doof* is not in first position, and why Jane Austen’s *Sense and Sensibility* was translated as *Gefühl und Verstand*, reversing the semantic order of the English title.

So, the voiced/voiceless contrast is not only phonetically absent from final obstruents in German words, plosives being realized as voiceless aspirated in citation form pronunciation, but there are also unaspirated/voiced realizations in fixed syntagmatic frames with high cohesion where rhythmicity becomes the binding principle. This syntagmatic domain of voicing exponency of German word-final obstruents as

part of speech rhythm needs systematic investigation in production and perception, and will not only advance our knowledge of segmental structures in phrase phonology but will also give us new insights into what constitutes speech rhythm.

Of course, incomplete neutralization would never have become an issue of phonetic analysis if American Structuralism had applied a polysystemic approach to phonology, as did Firth and his School [15, 16], instead of postulating an all-encompassing monolithic phoneme system, because then there would not have been anything to neutralize. The archiphoneme of the Prague Circle was already a step in this direction, and the extension of the issue of voicing exponency to the syntagmatic domain fits nicely into this polysystemic approach, providing another case that demonstrates the need to bridge the segment-prosody divide [17].

5. Epilogue

The different ways neutralization has been handled since the days of the Prague Circle may be given melodic expression.

- a Trubetzkoy made a phonological statement
There's neutrali'zation.
shown in Figure 2a: low pre-head, followed by an F0 rise-fall spread over the stressed and the following unstressed syllable, masked during the voiceless fricative;
- b Haskins phoneticians queried its phonetic adequacy
Neutrali'zation?
shown in Figure 2b: low pre-head, followed by a high F0 rise from the stressed-syllable onset to the end of the following unstressed syllable, masked during the voiceless fricative;
- c Dinnsen and the Laboratory Phonology community rejected it with phonological incredulity and indignation “Lady Bracknell” style [18]
Neutrali'zation?!
shown in Figure 2c: high pre-head, followed by a late high-rising valley contour with breathy intensification of the stressed syllable, masking F0 up to the sonorant of the unstressed syllable;
- d Wiktor Jassem brought it back to the phonetic level of American Structuralism with politely contrastive but compelling reference to phonetic reality
There's neutrali'zation!
shown in Figure 2d: low pre-head, followed by an F0 rise-fall-rise on the stressed and the following unstressed syllable, the fall being masked by the voiceless fricative;
- e I move the discussion from the paradigmatic monosystemic to the syntagmatic polysystemic level in a final conclusion
There's 'nothing to 'neutralize.
shown in Figure 2e: low pre-head, followed by an F0 rise in the first stress foot ‘nothing to and a subsequent early low F0 fall in the second stress foot ‘neutralize.

The prosodic embedding of segmental neutralization in this paper is a small personal tribute to Wiktor Jassem on the occasion of his 90th birthday in appreciation

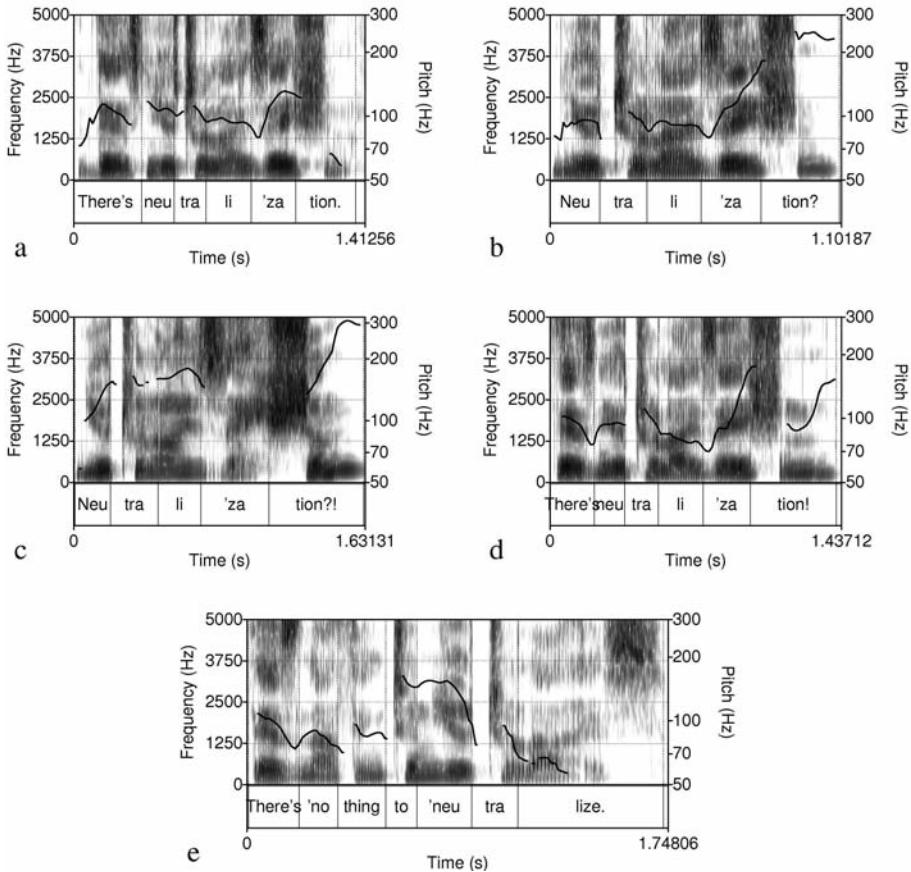


Figure 2: Spectrograms, F0 curves and syllabic segmentation of a “There’s neutralization.” (statement), b “Neutralization?” (confirmation question), c “Neutralization?!” (emphatic contradictory question), d “There’s neutralization!” (contrastive statement), e “There’s nothing to neutralize.” (matter-of fact conclusive statement). Speaker: the author.

of his broad spectrum of contributions to phonetic science from segments to rhythm and intonation.

REFERENCES

- [1] Jassem, W. and L. Richter. 1989. Neutralization of voicing in Polish obstruents. *Journal of Phonetics* 17, 317–325.
- [2] Kohler, K. J. 1995. *Einführung in die Phonetik des Deutschen*. Berlin: Erich Schmidt Verlag (2nd ed.).
- [3] Trubetzkoy, N. S. 1939. *Grundzüge der Phonologie*. 3rd ed. 1962 Göttingen: Vandenhoeck & Ruprecht.
- [4] Mitleb, F. 1981. *Segmental and Non-segmental Structure in Phonetics: Evidence from Foreign Accent*. Ph.D. dissertation, Indiana University, Bloomington.

- [5] Port, R., F. Mitleb, and M. O'Dell. 1981. Neutralization of obstruent voicing in German is incomplete. *Journal of the Acoustical Society of America* 70, S13, F10.
- [6] O'Dell, M., and R. Port. 1983. Discrimination of word-final voicing in German. *Journal of the Acoustical Society of America* 73, S31, N17.
- [7] Port, R., and M. O'Dell. 1985. Neutralization of syllable-final voicing in German. *Journal of Phonetics* 13, 455–471.
- [8] Fourakis, M., and G. K. Iverson. 1984. On the 'incomplete neutralization' of German final obstruents. *Phonetica* 41, 140–149.
- [9] Slowiaczek, L. M. and D. A. Dinnsen. 1985. On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13, 325–341.
- [10] Dinnsen, D. A. 1983. *On the Characterization of Phonological Neutralization*. Bloomington: Indiana University Linguistics Club.
- [11] Dinnsen, D. A. 1985. Re-Examination of Phonological Neutralization. *Journal of Linguistics* 21, 265–279.
- [12] Kohler, K.J. 2007. Beyond laboratory phonology: The phonetics of speech communication. In M. J. Solé, P. S. Beddor, M. Ohala, eds., *Experimental approaches to phonology*. Oxford: Oxford University Press. 51–53.
- [13] Röttger, T. B., Winter, B., and Grawunder, S. 2011. The robustness of incomplete neutralization in German. *Proc. 17th ICPhS, Hong Kong*. 1722–1725.
- [14] Kohler, K.J. 2009. Rhythm in speech and language. A new research paradigm. *Phonetica* 66, 20–45.
- [15] Firth, J.R. 1948. Sounds and prosodies. *Transactions of the Philological Society* 1948, 127–152.
- [16] Hawkins, S.; Smith, R. 2001. Polysp: a polysystemic, phonetically- rich approach to speech understanding. *Ital. J. Ling.* 13, 99–188.
- [17] K. Kohler, K. 2011. Communicative functions integrate segments in prosodies and prosodies in segments. *Phonetica* 68, 26–56.
- [18] Kohler, K.J. 2011. On the interdependence of sounds and prosodies in communicative functions. *Proc. 17th ICPhS, Hong Kong*. 19–27.