Grazer Linguistische Studien 76 (Herbst 2011); S. 55-74

The Nature of Incomplete Neutralization in German: Implications for Laboratory Phonology

Bodo Winter
University of California, Merced, Department of Cognitive and Information Sciences

Timo Röttger
IfL Phonetik, University of Cologne, Germany

Abstract. A number of researchers have contested the completeness of several neutralization phenomena, including final devoicing in German. These ideas received a fair amount of resistance because of methodological and theoretical concerns. In particular, it seemed that the research on Incomplete Neutralization involved the acceptance of certain assumptions from formal phonology such as the existence of underlying phonemes. In this paper, we discuss the evidence for Incomplete Neutralization and we show that no phonology-induced phantoms need to be invoked in order to explain these findings. Rather, there are multiple hypotheses that allow grounding Incomplete Neutralization in well-known experimental phenomena such as lexical co-activation and phonetic convergence. These accounts allow marrying phonological completeness with phonetic incompleteness, and they are consistent with a large body of experimental work from other disciplines.

Keywords. Incomplete neutralization, final devoicing, mental lexicon, laboratory phonology

1. Introduction

The contrast between German voiced and voiceless obstruents is generally thought to be lost in syllable-final position. For example, the plural words *Räder* 'wheels/bicycles' and *Räte* 'councils' are both pronounced [ʁaːt] in the singular. This distributional fact about German phontactics has been called final devoicing[1] and has been considered "the prime example of a phonological rule of Modern Standard German" (Wiese, 1996: 204), or even the "most popular of German phonological rules" (Giegerich, 1989: 51).

Final devoicing is often considered to be a purely phonological phenomenon. Theories such as Generative Phonology and Optimality Theory predict that there are no phonetic differences between a voiceless consonant such as /t/ in *Rat* and a

[1] While "tense" and "lenis" might be more accurate terms to describe German consonants (Kohler, 1984), we stick with "voiced" and "voiceless" to remain consistent with the literature on Incomplete Neutralization.

"devoiced" /d/ in *Rad*. This is because final devoicing is thought to happen in the phonological system prior to articulatory execution. The speech production system only receives this voiceless /t/ - it cannot "see" that the original stop was voiced in the lexicon. The resulting segment is completely voiceless.

Numerous researchers have contested this view. Based on phonetic analyses, it was argued that there are indeed acoustic differences between words such as *Rad* and *Rat*, and that these differences are perceivable (e.g. Dinnsen & Garcia-Zamor, 1971; Mitleb, 1981; Port et al., 1981; O'Dell & Port, 1983; Taylor, 1975; Port & O'Dell, 1984; Port et al., 1984; Charles-Luce, 1985; Dinnsen, 1985; Port & O'Dell, 1985; Port & Crawford, 1989; Piroth & Janker, 2004; Kleber et al., 2010; Röttger et al., 2011). These findings were taken to suggest a "semicontrast", a phenomenon that came to be labeled Incomplete Neutralization (IN).

Proponents of IN made strong claims such as: "German apparently does not have an abstract phonological rule of neutralization, despite almost a hundred years of assertions by linguists and German pedagogists that it does" (Port & Crawford, 1989: 280). Port and Crawford (1989: 257) furthermore describe IN as posing "a threat to phonological theory". Others extended these findings to all languages, proposing that "many putative neutralizations, when examined more carefully, may be shown to be non-neutralizing" (Dinnsen, 1985: 277). Port and O'Dell (1985: 466) even stated that researchers should consider the "radical hypothesis" that true neutralization never occurs.

IN effects of final devoicing have subsequently been found in Dutch, Catalan, Polish and Russian (Warner et al., 2004; Charles-Luce & Dinnsen, 1987; Slowiaczek & Dinnsen, 1985; Dmitrieva et al., 2010). However, the phenomenon is not restricted to final devoicing. Acoustic studies have revealed IN for spirantization in Eastern Andalusian Spanish (Gerfen & Hall, 2001; Gerfen, 2002; Bishop, 2007), consonant deletion in Turkish (Dinnsen, 1985), liquid neutralization in Puerto Rican Spanish (Simonet et al., 2008) and flapping of intervocalic alveolar stops in American English (de Jong, 2011).

Incomplete Neutralization is closely related to the phenomenon of near mergers. With near mergers, speakers report that two phonological categories are the same, but consistently differentiate them in production. Near mergers have been reported for vowels in many English dialects (Labov, 1971; Trudgill, 1974; Nunberg, 1980; Harris, 1985; Di Paolo, 1988). For example, Labov et al. (1972: ch. 6) demonstrated that New York speakers differentiate *source* and *sauce* in production but report no perceptual distinction between them. Near mergers are not restricted to segmental contrasts: Yu (2007) demonstrated that derived mid-rising tones in Cantonese show small but statistically significant differences from underived mid-rising tones.

The difference between Incomplete Neutralization and near mergers is primarily one that has to do with the history of investigation of that contrast (Manaster Ramer, 1996). In the case of Incomplete Neutralization, the majority of researchers thought the contrast to be neutralized, but then production studies found acoustic differences. In the case of near mergers, no such prior assumption was made, and often there was a recent sound change that resulted in the near merger. There is no difference in the

nature of the evidence that is taken to support IN, and the evidence that is taken to support near mergers.

Going back to German, the claim that final devoicing is incomplete stands against a long tradition, ranging from Jespersen (1913) and Trubetzkoy (1939) to modern descriptions such as Wiese (1996) and Zifonun et al. (1997), including discussions of final devoicing in introductory textbooks (e.g. Hyman, 1975; Lass, 1984; Hall, 2000). Moreover, IN is difficult to incorporate into formal theories of phonology, which generally have problems incorporating "in-between" categories (Port & Leary, 2005). Given the apparent inconsistency with received wisdom, it comes as no surprise that the idea of IN was heavily criticized. Some researchers tried to explain away the experimental results by pointing out methodological errors (Fourakis & Iverson, 1984; Manaster Ramer, 1996; Kohler, 2007). Fuchs (2005: 25) rightly says that the IN debate "became more and more a debate about methodological problems". The resulting confusion about the empirical evidence has led many researchers to acquire a firm disbelief in the existence of IN. Kohler (2007) furthermore argues that the idea is based on the faulty program of trying to find phonetic evidence for entities from formal phonology such as "underlying" voiced phonemes.

In this paper, we argue that IN is not a phonology-induced phantom. We discuss evidence that suggests that the phenomenon might actually have nothing to do with considerations of formal phonological theory at all. The terminology and conceptual framework of previous studies on IN were embedded in the framework of Generative Phonology, talking of underlying phonemes that do or do not surface in a neutralizing context. We believe that this has led to a lot of unnecessary resistance to the idea of IN, as acceptance of the phenomenon seemed to entail the existence of phonological categories such as phonemes and phonological processes such as "devoicing". We will show that we can talk about IN without this theoretical baggage by grounding the phenomenon in general cognitive and phonetic principles.

We first discuss the experimental evidence for IN in section 2, and the criticism of this evidence in section 3. In section 4, we compare traditional theoretical accounts with modern alternative accounts that have received less attention. We provide experimental evidence for these accounts and raise the possibility that IN might in fact have different mutually compatible explanations, none of which invoke formal phonological theory.

2. The empirical basis of Incomplete Neutralization

Production studies on IN often asked participants to read out minimal pairs such as *Tod* 'death' and *tot* 'dead', where small but significant acoustic differences between such words were found (Port et al., 1981; O'Dell & Port, 1983; Charles-Luce, 1985; Port & O'Dell, 1985)[2]. The most important correlates of voicing were the duration of

[2] Some studies also investigated articulatory factors (Fuchs & Perrier, 2003; Fuchs, 2005).

the preceding vowel, the closure duration, the duration of the "voicing-into-closure", as well as burst and aspiration duration. Among these correlates, the duration of the preceding vowel stands out as the most reliable one across studies and across languages. Importantly, in German these correlates are known to distinguish voiceless from voiced obstruents in intervocalic contexts (e.g., Keating, 1984; Kohler, 1984), and the direction of the differences resembles the non-neutralizing contrast.

An important property of the differences in IN studies is that they are generally very small, e.g. 1.2-6.2ms in German (Port & Crawford, 1989) or 3.5ms in Dutch (Warner et al., 2004). Moreover, the magnitude of the IN effect appears to be dialect- and speaker-dependent (Piroth & Janker, 2004), as well as highly sensitive to phonetic, semantic, and pragmatic context (Charles-Luce, 1985, 1993; Ernestus & Baayen, 2006; Port & Crawford, 1989; Slowiaczek & Dinnsen, 1985).

Moving on to perception studies of IN, listeners are usually fairly inaccurate at perceiving the final contrast, with people performing barely above chance level (see Brockhaus, 1995: 244 for an overview). As opposed to perceiving the voicing contrast in the initial or medial position, perception in the final position does not seem to be categorical (see Kleber et al., 2010 for German; Warner et al., 2004 for Dutch).

Brockhaus (1995: 244), among many others, points out that it is not clear whether the perceptual difference between syllable-final voiced and voiceless obstruents is actually "salient enough to be relied upon in normal communication". Although it is not known how accurate a contrast needs to be perceived in order to play a role outside the laboratory (Xu, 2010: 334), the low accuracy scores and the large variability suggest that IN has not much functional relevance in everyday communicative situations.

This is especially the case because in German, IN bears no functional load: the members of the minimal pair *Rad* 'wheel' and *Rat* 'council' would almost never occur in the same context for semantic reasons, and they also differ in grammatical gender, which means that the preceding article disambiguates the words. In German, almost any of the relevant minimal pairs is either distinguished by word class (e.g. *Tod* 'death' vs. *tot* 'dead') or by gender. Proponents of IN, too, recognize that the "semicontrast" has next-to-no functional load in German (Port & Crawford, 1989: 260; see also Fuchs, 2005: 173)[3].

The likely absence of any functional relevance might suggest to some that the phenomenon is not worth studying. However, after discussing some methodological problems with IN studies in the next section, we will point out that even small differences that might actually play no role in everyday speech communication can be of importance for theoretical reasons: even small effects can give us hints at the cognitive architecture that is at work when people use spoken language.

3. Methodological problems

In this section, we will briefly discuss the many concerns that have been raised against IN. It is important to review the evidence methodologically before we address

[3] Note that in Dutch, IN might signal past tense (Ernestus & Baayen, 2006).

possible causes of IN – if the results are spurious there is no need to come up with explanatory accounts.

3.1. Orthography

In German and some other languages, the difference between voiced and voiceless categories is orthographically preserved (e.g. *Rad* vs. *Rat*). Opponents argue that this orthographical distinction is the sole cause of IN, where participants are thought to perform "artificial" hypercorrection based on the written language. In early studies, the analyzed utterances were based on read speech, making orthography particularly salient. When a completely auditory task design was used, results were mixed, e.g. Fourakis and Iverson (1984) found no IN effect, but Port and Crawford (1989) did.

Both of these auditory studies, however, are prone to methodological criticism: first, Fourakis and Iverson (1984) tried to discount IN, but they used a very small sample of speakers, inviting the criticism that the statistical power was too low to detect potential IN effects[4]. However, the numerical differences in this study went into the right direction (mirroring the non-neutralized contrast), and a re-analysis by Port and Crawford (1989: 259) found statistically reliable effects. Many studies interpret Fourakis and Iverson (1984) as evidence against the existence of IN (Jessen, 1998: 335; Kopkalli, 1993; Kohler, 2007: 45; Manaster Ramer, 1996: 481; Wiese 1996: 205), however, with such a small sample size and the mentioned re-analysis, this does not seem to be warranted[5].

Second, Port and Crawford (1989) tried to completely eliminate the role of orthography: they read the critical stimuli to the participants who had to repeat them. Given that we know that participants accommodate to the experimenter (Hay et al., 2009), the influence of orthography might just have carried over from the experimenter to the participant. Thus, while many researchers interpret Port and Crawford (1989) as successfully eliminating the role of orthography (e.g. Kopkalli, 1993: 130), this is clearly not the case.

Röttger et al. (2011) addressed the issue of orthography by using a design that was completely auditory, where participants were presented spoken pseudowords with a voiced or voiceless obstruent in a non-neutralized position, such as in *die Gobe*. Participants then had to produce a morphologically related form where the segment appeared in a neutralized position: *ein Gob*. Given that the participants never saw these pseudowords in their written form, the effect of orthography in this design was minimized, yet, a robust IN effect for vowel duration was found. Röttger and colleagues also had more speakers than any previous study on German IN, addressing the above-mentioned concern of statistical power.

[4] Frick (1995) argues that one can only gain confidence in a null hypothesis (e.g. "there is no Incomplete Neutralization") if one demonstrates "sufficient effort" to disprove the null. Given that Fourakis and Iverson (1984) had fewer subjects and items than comparable studies, one cannot take this study to show that IN does not exist. At a bare minimum, studies that try to discount IN have to have at least as many subjects as studies supporting the phenomenon.

[5] The same reasoning applies to other IN studies with small samples, such as Jassem and Richter (1989) for Polish, and Kopkalli (1993) for Turkish.

Does this mean that orthography plays no role at all? We have several strands of evidence suggesting that it does. O'Dell and Port (1983), Port and O'Dell (1985), and Port and Crawford (1989: 271) found differences between such words as *seid* and *seit*. As there are no relevant morphological alternations for *seid* which permit the conclusion that there is a /d/ in the final position, it seems that the acoustic differences are most likely due to orthography (Manaster Ramer, 1996; Kohler, 2007). Warner and colleagues (2006) and Ernestus and Baayen (2006) more conclusively showed the influence of orthography in Dutch by finding IN effects for a contrast that was only orthographically represented. Warner et al. (2004: 253) furthermore argued that across different languages, larger differences are found for those contrasts that are orthographically represented. And, experimental designs that emphasize the writing system obtain stronger IN effects (Port & Crawford, 1989).

Ultimately, the problem of orthography is not about experimental items or designs, but about the investigated speaker population. As adult literate speakers habitually associate speech with orthographic representations, any experiment that tests this population cannot completely rule out the influence of the written language. So why have proponents of IN not tested illiterate speaker populations? There are several reasons: first, illiteracy is highly variegated, and a lot of illiterates have had and continue to have exposure to writing. 'True' illiterates who have never received any schooling are difficult to find in the relevant languages. Second, in developed countries, illiterates are always surrounded by literates which – if they produce slight differences due to their orthographic representations – might influence illiterates through the forces of phonetic convergence. This ultimately means that the role of orthography will probably always remain an issue with respect to IN.

What is the evidence that IN does not completely depend on orthography? Catalan (Dinnsen & Charles-Luce, 1984) shows IN, and Turkish demonstrates numerical trends of IN (Kopkalli, 1993), even though both languages do not make the relevant orthographic distinctions. So, we can say that orthography clearly plays a role, but IN can occur regardless of it.

3.2. Infrequent and obsolete words

One common criticism of IN experiments is the choice of experimental stimuli presented to the participants (Manaster Ramer, 1996; Kohler, 2007). For example, Kohler (2007: 45) found it problematic that for some IN studies, researchers had to explain the meanings of some of their words to their participants before the experiment. We agree with this concern.

There are three problems associated with using infrequent or obsolete items. One harkens back to our discussion of orthography. If participants have no knowledge of the word, they might turn to orthography to base their phonological decisions (e.g. is this a /d/ or /t/?). The second problem is that these items might draw attention to the experimental manipulation (Kopkalli, 1993: 7-8). While this is a concern in many of the older studies, the relatively large amount of filler items in

other studies (Piroth & Janker, 2004; Port & Crawford, 1989; Röttger et al., 2011) makes it unlikely that the obviousness of the task plays a huge role.

The final problem is that we know that very infrequent forms tend to be hyperarticulated with respect to very frequent forms that tend to be shortened (e.g. Aylett & Turk, 2004). Hyperarticulated word forms are more likely to preserve phonemic contrasts, in particular duration contrasts, and therefore, the infrequent and obsolete words might have made it much easier for IN effects to reach statistical significance.

However, we argue that the ease of finding statistical effects is the only concern with respect to hyperarticulation. It does not completely discount IN because hyperarticulation assumes the presence of IN: there needs to be some phonetic or phonological material that is present before becoming hyperarticulated. Moreover Röttger et al. (under review) found IN to be unaffected by prosodic position – even though one would expect effects to be stronger in a prominent prosodic position if hyperarticulation played a role.

Related to hyperarticulation is the idea that participants try to distinguish homophones in experiments on IN. Charles-Luce (1993) showed that IN effects in Catalan disappear when the semantic context disambiguated the lexical form. But the idea that participants are (consciously or non-consciously) enhancing a contrast to distinguish homophones is again no knock-down argument. The same reasoning as with hyperarticulation applies: the increased difference when distinguishing two homophones must be based on a contrast that has to be there to begin with. Furthermore, if homophony-distinguishing behavior were the sole cause of IN, we would expect the differences to fade away in designs with many fillers, where "minimal pairs" (e.g. the "homophones" *Rad* and *Rat*) are less obvious. This prediction is not born out by the evidence (Dinnsen & Charles-Luce, 1984 for Catalan; Piroth & Janker, 2004; Röttger et al., 2011 for German).

3.3. The influence of English

Kohler (2007) points out that all German-speaking populations tested so far had proficiency in English, which is a potential problem because English preserves the final voicing contrast (e.g. *bad* vs. *bat*, *bed* vs. *bet*). Early studies on IN conducted their experiments with German speakers in English-speaking environments or with English-speaking experimenters present (Mitleb, 1981; O'Dell & Port, 1983; Port et al., 1983; Port & O'Dell, 1985; Port & Crawford, 1989). This is problematic because we know that the experimenter and the testing environment can bias the results of phonetic experiments (Hay et al., 2009; Hay & Drager, 2010). While more recent studies tested German native speakers in a German-speaking environment (Fuchs et al., 2005; Kleber et al., 2010; Piroth & Janker, 2004; Röttger et al., 2011), the influence of English as a second language is still a concern because we know that L2 categories can influence L1 categories on a long-term basis (e.g. Chang, 2011). Again, to resolve this issue, the evidence from Cantonese and Catalan is crucial (Yu, 2007; Dinnsen & Charles-Luce, 1984): in these languages, there are no obvious L2 influences that play a role with respect to the investigated contrasts. We therefore predict that IN effects

should also be found with German speakers with only minor or no exposure to English.

3.4. Summing up: converging evidence?

The acceptance of any phenomenon should never be based on a single study and several studies, such as Fourakis and Iverson (1984), have been overemphasized in their importance relative to the totality of IN studies. Only by accumulating lots of converging evidence from different methodologies can we be more certain about the existence of IN. It seems that by now, there are far more studies finding evidence of IN (both within German and across languages) than finding counter-evidence. Richard Wiese (1996: 205) commented on IN experiments as follows: "These results are rather tentative […] given that the recognition of non-neutralized devoicing was found in a minority of cases only". By now, we can – with relative confidence – say that this statement turned out not to be true. Positive results for IN characterize the majority of studies on this topic and several of the methodological issues have been successfully addressed. Now that we have discussed the evidence for IN, we need to turn to what might cause this phenomenon.

4. The nature of Incomplete Neutralization

4.1. IN and formal phonology

Most early studies supporting IN couched their findings in the conceptual language of Generative Phonology. In this section, we will give a quick overview of the kinds of theoretical proposals that have been made. For example, Port and O'Dell (1985: 466) discuss the possibility of a feature [voice-F] that applies to the special case of IN, implemented in between the voiced and voiceless category. Charles-Luce (1985: 319-323) argues for final devoicing being a feature-deletion rule, rather than a feature-changing rule where [+obstruent] becomes [-voice]. The resulting segment would be unspecified for voice but marked for having undergone the final devoicing process (cf. Piroth & Janker, 2004, who tentatively argue for viewing final stops as unspecified archiphonemes).

Port and O'Dell (1985: 468-469) finally argue for a phonetic implementation rule that devoices at the syllable level. The phonological [+voice] feature that is assumed to be in the lexicon is left unchanged, but phonetically, and only at the level of the syllable, a [-voice] gesture is realized. This is similar to Fuchs (2005: 173), who links neutralization to a process of articulatory reduction, allowing for speaker-specific "articulatory residues of the contrast". Another approach is to ascribe the IN effect to rule ordering (Charles-Luce, 1985: 319). Here, the phonological devoicing rule is ordered after or simultaneously with a phonetic implementation rule. This means that the [voice] feature is still available to the implementation and is therefore allowed to surface.

Within the framework of Government Phonology, Brockhaus (1995: 250-251) suggests that the acoustic difference between a truly voiceless segment and a

devoiced one can be explained by a representational difference, where the originally voiced segments (which were employed with a laryngeal element L for "slack vocal folds") have a non-licensed L after undergoing final devoicing. Truly voiceless segments, on the other hand, have an H element for "stiff vocal folds".

In an attempt to model IN within the framework of Optimality Theory, van Oostendorp (2008) argues that output structures can be characterized through projection and pronunciation relations. Projection relations are abstract relationships between segments and their features. Pronunciation relations are output relationships between the feature and the segment, describing the actual phonetic output. This approach leads to three different categories: segments that are underlyingly voiced and pronounced voiced (e.g. [d] in *Räder*), segments that are underlyingly voiceless and pronounced voiceless (e.g. [t] in *Räte*), and segments that are underlyingly voiced but pronounced as voiceless (e.g. [t] in *Rad*). The latter type is allowed to be phonetically different from a truly voiceless segment because it is characterized by a different structure in the pronunciation relations.

These discussions, more than anything, showcase how the debate about IN was (and occasionally still is) heavily influenced by considerations of formal phonology. In discussing the IN findings, authors saw their results as standing against certain theories, and they assumed that their results required modifications to these theories. Kohler (2007) characterizes the early IN literature as "phonology-going-into-the-lab", where researchers simply try to confirm or disconfirm phonological concepts and theories via phonetic experiments. Most of the phonological proposals that were based on this approach lead to a proliferation of features or rules. And, the proposals were predominantly post-hoc and had no predictive power.

To us, most formal phonological accounts of IN seem like rescue attempts that try to make phonological theory fit to the experimental facts, or that try to fit the data into existing theories. This merely post-dicts the data and does not lead to new testable hypotheses. We agree that in principle, IN is of "tremendous phonological interest" (Fourakis & Iverson, 1984: 141), but we do not think that the phenomenon is necessarily connected to formal phonology in the way it was discussed in the literature. We think that there potentially is a much more complex and interesting picture, where IN can be seen as emerging from general cognitive and phonetic processes. In fact, we argue that IN can be explained even without recourse to the phoneme, making almost no theoretical assumptions.

In section 4.2 we discuss the possibility of IN springing out of spreading activation in the mental lexicon. In section 4.3 we discuss the role of phonetic convergence in some experiments on IN. We propose that next to orthography, both of these accounts might play a role in explaining IN.

4.2. Spreading activation

The mental lexicon is assumed to contain representations for a large number of words, and evidence suggests that fully inflected forms are stored as well (Alegre & Gordon, 1999; Baayen et al., 1997; Bybee, 1995; Sereno & Jongman, 1997). In this model of lexical organization and access, German speakers would have inflected

forms such as *Räder* 'wheels' in their mental lexicon. Due to its morphological relation with the singular form *Rad*, these two forms will be closely connected to each other in the relational network of the lexicon. Ernestus and Baayen (2006) consider the possibility that IN results from co-activation of morphologically related forms, i.e. when speakers pronounce *Rad*, they also activate the non-neutralized *Räder* and *Rades* (genitive). If some or most of the co-activated forms contain a non-neutralized segment that is fully voiced, these voiced forms could influence the motor commands used in speech production in subtle ways, leading to the observed IN effects.

This approach is based on spreading activation, a phenomenon that has tremendous experimental support (e.g., see references in Collins & Loftus, 1975). More concretely with respect to morphological relations, there is evidence for the automatic activation of morphological neighbors (e.g. Andrews, 1989; Sears et al., 1995), as well as evidence for the influence of neighborhood density on speech production (Wright, 2004; Munson & Solomon, 2004; Munson, 2007), where words with more neighbors tend to be hyperarticulated.

If the morphological neighbors have a non-neutralized segment, then this phonetic detail might leak through into speech production because of spreading activation. This is the only extra assumption that this account of IN has to make: that the phonetic content of the co-activated representation affects the production of the target. Everything else does not require any extra assumptions because we know from independent experiments that spreading activation and co-activation of morphological neighbors occurs. This renders this account relatively parsimonious.

The co-activation hypothesis has the advantage of making testable predictions. First, the frequency of morphological neighbors should modulate the size of IN effects. A word that has highly frequent morphological neighbors with non-neutralized obstruents should exhibit more Incomplete Neutralization than a word that has less frequent morphological neighbors. For example, *Räder* and *Rades* are relatively frequent with respect to *Rad*, whereas *die Braven* 'the well-behaved ones' and *die braven X* 'the well-behaved X' are relatively more infrequent compared to the frequent base form *brav* 'well-behaved'. Therefore, IN effects should be stronger for *Rad* than for *brav*. Another prediction is that there should be priming effects, for example, priming *Räder* should increase the partial voicing of *Rad*. And, the time course of the priming should modulate the effect: priming a few milliseconds before the production of *Rad* should lead to stronger effects than priming a few minutes before.

Now, it is not clear whether the spreading activation necessarily has to come from morphologically related forms or whether it might come from somewhere else. An alternative approach could be based on more global knowledge about phonotactic asymmetries. For example, in German, phonologically short vowels are more likely to precede voiceless obstruents. We know that listeners adjust to frequently occurring linguistic patterns (e.g. Hay et al., 2003; Pitt & McQueen, 1998), and it has been demonstrated that listeners are sensitive to phonotactic probabilities that are relevant for IN (Kleber et al., 2010). Spreading activation to phonotactic attractors in the brain cannot explain all IN effects, only the ones that involve vowel duration. But,

phonotactic probabilities might be one of the reasons why vowel duration is among the most robust IN cues.

4.3. Phonetic convergence

Convergence is another possible source of IN, however only in those experiments where participants have to repeat or respond to auditory stimuli (e.g. Port & O'Dell, 1989; Röttger et al., 2011). We know that people converge automatically and rapidly to the speech of others (Goldinger, 1998; Nielsen, 2005; Pardo, 2006) regardless of whether they like or do not like their interlocutor (Staum Casasanto et al., 2010).

As mentioned above, in Röttger et al. (2011), participants had to transform words such as *Gobe* into words such as *Gob*. There was minimal delay between the presentation of the auditory stimulus and the signal for presentation (500ms), which means that the just-heard word was still very active in memory, allowing for convergence to the auditory stimulus *Gobe*. A post-hoc analysis showed that the vowel duration of the intervocalic stimulus was a strong predictor for the vowel duration of the produced singular form, suggesting the influence of convergence. The vowel duration effect is clearly – at least to some extent – caused by phonetic convergence to the heard stimuli, but it looks just like Incomplete Neutralization in other experiments. In a follow-up study, we found IN regardless of convergence (Röttger et al., 2012), suggesting that while convergence plays a role, it cannot be the sole cause of the results in Röttger et al. (2011).

The phonetic convergence account also makes testable predictions. As convergence is modulated by sympathy for the speaker (Staum Casasanto et al., 2010), there should be more IN if speakers are more sympathetic to specific voices. Moreover, because we know that convergence is stronger for low frequency items than for high frequency ones (Goldinger, 1998), IN effects in a task such as the one used in Röttger et al. (2011) should be stronger for low frequency words.

5. Discussion

5.1. General discussion

In the last section, we have outlined two different accounts (co-activation and phonetic convergence) that likely play a role in causing IN. These accounts work in conjunction with orthography-induced IN. We view these three accounts as mutually compatible with each other, and we think that IN – just like many other phenomena – might have multicausal origins. At first sight, it seems unparsimonious to assume the truth of three different explanatory accounts for the same phenomenon. But, we know that morphological co-activation occurs, we know that people converge to the speech of others, and we know that orthographic forms become automatically activated when a word is processed (Perre et al., 2009; Seidenberg & Tanenhaus, 1979; Ziegler & Ferrand, 1998). The question then rather becomes: why should IN not result from these cognitive processes?

Our ideas are couched in explanatory pluralism (Dale et al., 2009; Mitchell, 2004; Mitchell & Dietrich, 2006), an approach to philosophy of science which recognizes the possibility that multiple hypotheses are equally valid and can equally help to explain a given phenomenon. In this paper, we have tried to synthesize different strands of research, showing how they can all be brought to bear on IN. This allows for formulating richer hypotheses for future studies. Rather than trying to disprove a given hypothesis or showing that one is primary, we argue that it is useful to consider the inter-relationships between these different hypotheses.

A pluralistic view of IN also acknowledges that formal phonology is not incompatible with our accounts. For example, one could still believe in abstract encapsulated phonological entities that are categorical, and one could still believe in complete neutralization within a phonological "module" – but in this perspective, the messiness of IN can be relegated to speech production, which is seen as a "hungry" process that takes every information that it can get (e.g. from orthographic representations and morphological neighbors). This view allows marrying phonological completeness with phonetic incompleteness.

However, we should point out that all three different accounts of IN outlined above are theoretically more in line with distributed processing accounts and experimentally based phonological theories such as exemplar-based theories. While we showed that we can talk about IN without even evoking the notion of a phoneme (let alone any features, rules or constraints), all our hypotheses point to a speech production system that is heavily distributed and interaction-dominant. For example, the orthographical explanation necessitates that speech production "sees" orthographical representations (Manaster Ramer, 1996), and the co-activation hypothesis entails that speech production "sees" morphological neighbors, and that the phonetics from these neighbors "leak into" a given utterance. And, phonetic convergence generally entails that categories are malleable and highly dependent on social interaction.

It might be that speech production is the messy and continuous cognitive system, whereas phonology remains "pure" and encapsulated (for the discussion of a similar argument with respect to the sensorimotor systems and higher-order cognition, see Spivey, 2007). But, to us, it seems rather unparsimonious to assume the existence of two different cognitive architectures, one distributed, another one encapsulated. We see the burden of proof resting with formal phonological theories (and abstract symbolic approaches more generally) that assume binary categories to show how a system that we know is physically instantiated via distributed population codes[6] and interactions between subsystems can realize abstract clear-cut categories.

Moreover, accounts that do not make recourse to formal phonology have the added advantage of making experimentally testable predictions (e.g. priming effects of morphological neighbors). It is not clear to us how phonological approaches to IN (including recent ones such as van Oostendorp, 2008) go beyond fitting the data into

[6] For example, Averbek et al. (2006) say that "individual neurons count for little; it is the population of activity that matters".

a theory, and how these approaches can be used to make new predictions for experiments.

5.2. The role of exemplar-based models

One theory that has conceptually been in the background of our discussion is the exemplar-based approach to phonology. We believe that this approach has considerable things to say about IN. It has been shown that lexical representations contain detailed phonetic information of individual word forms (Brown & McNeill, 1966; Goldinger, 1997; Palmeri et al., 1993; Pisoni, 1997), and this phonetic detail is exploited in perception and word recognition (Davis et al., 2002; Hawkins & Nguyen, 2003). This has led researchers to turn to models that assume storage of concrete phonetic events. In these models, categories are defined by clouds of memorized tokens (exemplars), and each new experience changes the entire category system slightly (Hintzman, 1986; Nosofsky, 1986; Goldinger, 1996). In these models, the targets of speech production result from averaging exemplars within a region of an exemplar cloud (see Pierrehumbert, 2001; Pierrehumbert, 2002).

It can immediately be seen that these exemplar-based models are very consistent with our explanatory accounts mentioned above. Inflected forms (e.g. *Räder*) are stored in exemplar clouds topologically close to, and overlapping with, the clouds of related forms (*Rad*). This allows for heavy interactions between morphologically related forms in speech production. The exemplar cloud way of describing the phenomenon is nothing but a state space description of the neural activation patterns that characterize the close connection between morphological neighbors.

However, rather than just being merely compatible with our proposals, exemplar-based theories offer an intriguing new perspective on IN. As speech production targets are averaged over exemplar clouds in these models, a few outliers in an exemplar cloud can bias the distribution and lead to small changes in phonetic detail. If, for example, a word such as *Rad* is produced only a few times in a hypercorrected fashion based on orthography, these hypercorrected forms enter the exemplar cloud for *Rad*, which can affect subsequent productions via the averaging process. And, because of phonetic convergence, listeners' exemplar clouds and their subsequent productions will also be affected by hearing hypercorrected forms of *Rad*. Through this process, IN can remain in the speech of a community, and the neutralization is always kept from being fully complete – there only need to be enough "outlier" exemplars to begin with. This would be a more long-term effect of orthography and hypercorrection more generally, rather than the short-term effect that is caused by experimental task demands (which is what is usually considered in the IN debate).

The general idea of this could be verified computationally, but it also produces testable experimental hypotheses, e.g. when participants are thinking actively about the formal written language (e.g. when dictating a text), there should be "artificial" differences between final voiced and voiceless stops that are stronger than regular IN effects. In subsequent productions, there should be more IN-like effects that are

characterized by relatively smaller differences, and if one conducts an iterated learning experiment (e.g. Galantucci, 2005) with vertical transmission between different sets of participants, traces of the initial hypercorrection should be detectable throughout participant generations. If this view can be supported by empirical evidence, then this means that IN does not have to be something that happens all the time – occasional hypercorrections and extreme productions will suffice to keep IN-like differences in the linguistic system.

6. Conclusions

Manaster Ramer (1996: 487) uses the IN debate as a call for phonologists and phoneticians to collaborate more with each other. He points out that phoneticians need phonologists to design experiments that are phonologically meaningful, and phonologists need phoneticians to constrain their theories with empirical data. In Manaster Ramer's words (ibid. 487), "Phonologists cannot afford to be neutral" with respect to IN. We have argued above that IN has vast implications for phonology. However, we have also shown that the phenomenon can be seen in a different light if psycholinguistic and cognitive evidence is taken into account. So, not only do phonologists and phoneticians need to collaborate with each other, but both need to look more at work from other disciplines to gain new perspectives of old phenomena.

In particular, we have shown that there are other possible accounts that are consistent with what we know about a heavily interconnected mental lexicon, as well as being consistent with what we know about phonetic convergence and the influence of orthography. The presence of these accounts and the already existing experimental support for these ideas means that one can study IN without invoking formal phonological notions. To talk about IN, we do not necessarily have to do what Kohler (2007) criticized as "phonology-going-into-the-lab".

If IN is seen in this new light, we can also see that the phenomenon might become an interesting test bed for ideas about the mental lexicon and the speech production system. We believe that IN can be used to probe into the cognitive mechanisms that underlie morphological co-activation and its influence on speech production, as well as to probe into the structure of the mental lexicon. Thus, IN could become elevated from being a "quirk" of German and some other languages to become an access point for cognitive investigations of speech.

Either way, our discussion above has emphasized that the evidence for IN is solid, and that rather than trying to question or discount the phenomenon, we now have to move towards studying its nature more deeply, and to work on collecting and synthesizing more evidence for different explanatory accounts. In this paper, we have provided a first attempt in this direction.

7. References

Alegre, Maria & Gordon, Peter. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, 40, 41-61.

Andrews, Sally (1989). Frequency and neighborhood effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 802-814.

Averbek, Bruno B. & Latham, Peter E. & Pouget, Alexandre (2006). Neural correlations, population coding and computation. *Nature Reviews Neuroscience*, 7, 358-366.

Aylett, Matthew & Turk, Alice (2004). The smooth signal redundancy hypothesis: a functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language & Speech*, 47, 31-56.

Baayen, R. Harald & Dijkstra, T. & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual route model. *Journal of Memory and Language*, 37, 94-117.

Bishop, Jason B. (2007). Incomplete neutralization in Eastern Andalusian Spanish: perceptual consequences of durational differences involved in s-aspiration. *Proceedings of the 16th ICPhS,* 1765-1768.

Brockhaus, Wiebke (1995). *Final Devoicing in the Phonology of German*. Tübingen: Max Niemeyer Verlag.

Brown, Roger & McNeill, David (1966). The 'tip of the tongue' phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 5, 325-337.

Bybee, Joan (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, 10, 425-455.

Chang, Charles B. (2011). Systemic drift of L1 vowels in novice L2 learners. *Proceedings of the 17th ICPhS,* 428-431.

Charles-Luce, Jan (1985). Word-final devoicing in German: effects of phonetic and sentential contexts. *Journal of Phonetics*, 13, 309-324.

Charles-Luce, Jan (1993). The effects of semantic context on voicing neutralization. *Phonetica*, 50, 28-43.

Charles-Luce, Jan & Dinnsen, Daniel A. (1987). A reanalysis of Catalan devoicing. *Journal of Phonetics*, 15, 187-190.

Collins, Allan M. & Loftus, Elizabeth F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407-428.

Dale, Rick & Dietrich, Eric & Chemero, Anthony (2009). Explanatory pluralism in cognitive science. *Cognitive Science*, 33, 739-742.

Davis, Matthew H. & Marslen-Wilson, William D. & Gaskell, M. Gareth (2002). Leading up the lexical garden-path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 218-244.

de Jong, Kenneth (2011). Flapping in American English. In M. van Oostendorp & C. J. Ewen & E. Hume & K. Rice (Hgg.), *Blackwell Companion to Phonology*. Oxford, UK: Wiley-Blackwell, pp. 2711-2729.

di Paolo, Marianna (1988). Pronunciation and categorization in sound change. In K. Ferrara & B. Brown & K. Walters & J. Baugh (Hgg.), Linguistic change and contact: *Proceedings of the 16th. Annual Conference on New Ways of Analyzing Variation in Language*. Austin: University of Texas, pp. 84-92.

Dinnsen, Daniel A. & Garcia-Zamor, Maria (1971). The three degrees of vowel length in German. *Papers in Linguistics*, 4, 111-126.

Dinnsen, Daniel A. (1985). A re-examination of phonological neutralization. *Journal of Linguistics*, 21, 265-279.

Dinnsen, Daniel A. & Charles-Luce, Jan (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics*, 12, 49-60.

Dmitrieva, Olga & Jongman, Allard & Sereno, Joan (2010). Phonological neutralization by native and non-native speakers: The case of Russian final devoicing. *Journal of Phonetics*, 38, 483-492.

Ernestus, Mirjam & Baayen, R. Harald (2006). The functionality of incomplete neutralization in Dutch: the case of past-tense formation. In L. M. Goldstein & D. H. Whalen & C. T. Best (Hgg.), *Laboratory Phonology VIII*. Berlin: Mouton de Gruyter, pp. 27-49.

Fourakis, Marios & Iverson, Greg (1984). One the 'Incomplete Neutralization' of German Final Obstruents. *Phonetica*, 41, 140-149.

Frick, Robert W. (1995). Accepting the null hypothesis. *Memory & Cognition*, 23, 132-138.

Fuchs, Susanne (2005). Articulatory correlates of the voicing contrast in alveolar obstruent production in German. *ZAS Papers in Linguistics*, 41.

Fuchs Susanne & Perrier, Pascal (2003). An EMMA/EPG study of voicing contrast correlates in German. *Proceedings of the 15th ICPhS*, 1057-1060.

Galantucci, Bruno (2005). An experimental study of the emergence of human communication systems. *Cognitive Science*, 29, 737-767.

Gerfen, Chip (2002). Andalusian Codas. *Probus*, 14, 247-277.

Gerfen, Chip & Hall, Kathleen (2001). Coda aspiration and incomplete neutralization in Eastern Andalusian Spanish. Manuscript, University of North Carolina at Chapel Hill. Retrieved from www.unc.edu/~gerfen/papers/GerfenandHall.pdf

Giegerich, Heinz J. (1989). *Syllable structure and lexical derivation in German*. Indiana University Linguistics Club.

Goldinger, Stephen D. (1996). Words and voices: Episodic traces in spoken word identification and recognition in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1166-1183.

Goldinger, Stephen D. (1997). Words and voices: Perception and production in an episodic lexicon. In K. Johnson & J. W. Mullennix (Hgg.), *Talker Variability in Speech Processing*. San Diego: Academic Press, pp. 33-65.

Goldinger, Stephen D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251-279.

Hall, T. Alan (2000). *Phonologie: Eine Einführung*. Berlin & New York: Mouton de Gruyter.

Harris, John (1985). *Phonological variation and change*. Cambridge: Cambridge University Press.

Hawkins, Sarah & Nguyen, Noël (2003). Effects on word recognition of syllable-onset cues to syllable-coda voicing. In J. Local & R. Ogden & R. Temple (Hgg.), *Papers in laboratory phonology 6* . Cambridge: Cambridge University Press, pp. 38-57.

Hay, Jennifer & Drager, Katie (2010). Stuffed toys and speech perception. *Linguistics*, 48, 865-892.

Hay, Jennifer & Drager, Katie & Warren, Paul (2009). Careful who you talk to: An effect of experimenter identity on the production of the NEAR/SQUARE merger in New Zealand English. *Australian Journal of Linguistics*, 29, 269-285.

Hay, Jennifer & Pierrehumbert, Janet & Beckman, Mary (2003). Speech perception, well-formedness and the statistics of the lexicon. In J. Local & R. Odgen, & R. Temple (Hgg.), *Papers in Laboratory Phonology VI*. Cambridge: Cambridge University Press, pp. 58-74.

Hintzman, Douglas L. (1986). Schema abstraction in a multiple-trace memory model. *Psychological Review*, 93, 411-428.

Hyman, Larry M. (1975). *Phonology: Theory and Analysis*. Fort Worth: Harcourt Brace Jovanovich College Publishers.

Jassem, Wiktor & Richter, Lutoslawa (1989). Neutralization of voicing in Polish obstruents. *Journal of Phonetics*, 17, 317-325.

Jespersen, Otto (1913). *Lehrbuch der Phonetik*. Leipzig: G.B. Teubner.

Jessen, Michael (1998). *Phonetics and Phonology of Tense and Lax Obstruents in German*. Amsterdam: John Benjamins.

Keating, Patricia A. (1984). Phonetic and Phonological Representation of Stop Consonant Voicing. *Language*, 60, 286-319.

Kleber, Felicitas & John, Tina & Harrington, Jonathan (2010). The implications for speech perception of incomplete neutralization of final devoicing in German. *Journal of Phonetics*, 38, 185-196.

Kohler, Klaus J. (1984). Phonetic Explanation in Phonology: The Feature Fortis/Lenis. *Phonetica*, 41, 150-174.

Kohler, Klaus J. (2007). Beyond Laboratory Phonology. The phonetics of speech communication. In M.-J. Solé & P. S. Beddor & M. Ohala (Hgg.), *Experimental Approaches to Phonology*. Oxford: Oxford University Press, pp. 41-53.

Kopkalli, Handan (1993). *A Phonetic and Phonological Analysis of Final Devoicing in Turkish*. University of Michigan, PhD Thesis.

Labov, William (1971). Methdology. In W. O. Dingwall (Hg.), *A survey of linguistic science*. College Park: University of Maryland Linguistics Program, pp. 412-497.

Labov, William & Yaeger, Malcah & Steiner, Richard (1972). *A quantitative study of sound change in progress*. Philadelphia: US Regional Survey.

Lass, Roger (1984). *Phonology: An introduction to basic concepts*. Cambridge: Cambridge University Press.

Mitchell, Sandra D. (2004). Why integrative pluralism? *E:CO*, 6, 81-91.

Mitchell, Sandra D. & Dietrich, Michael (2006). Integration without Unification: An Argument for Pluralism in the Biological Sciences. *The American Naturalist*, 168, 73-79.

Munson, Benjamin (2007). Lexical access, lexical representation, and vowel production. In J. S. Cole & J. I. Hualde (Hgg.), *Laboratory Phonology IX*. Berlin: Mouton de Gruyter, pp. 201-228.

Munson, Benjamin & Solomon, Nancy Pearl (2004). The effect of phonological density on vowel articulation. *Journal of Speech, Language and Hearing Research*, 47, 1048-1058.

Nielsen, Kuniko (2005). Specificity and Generalizability of Spontaneous Phonetic Imitation. *Proceedings of the 9th International Conference on Spoken Language Processing*. Pittsburgh.

Nosofsky, Robert M. (1986). Attention, similarity and the identification–categorization relationship. *Journal of Experimental Psychology: General*, 115, 39-57.

Nunberg, G. (1980). A falsely reported merger in eighteenth-century English: A study in diachronic variation. In W. Labov (Hg.), *Locating language in time and space*. New York: Academic Press, pp. 221-250.

O'Dell, Michael & Port, Robert F. (1983). Discrimination of word final voicing in German. *Journal of the Acoustical Society of America*, 73, suppl. 1.

Palmeri, Thomas J. & Goldinger, Steven D. & Pisoni, David B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 309-28.

Pardo, Jennifer S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119, 2382-2393.

Perre, Laetitia & Midgley, Katherine & Ziegler, Johannes C. (2009). When beef primes reef more than leaf: Orthographic information affects phonological priming in spoken word recognition. *Psychophysiology*, 46, 739-746.

Pierrehumbert, Janet B. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P. Hooper (Hgg.), *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins, pp. 137–158).

Pierrehumbert, Janet B. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (Hgg.), *Laboratory Phonology 7*. Berlin & New York: Mouton de Gruyter, pp. 101-139.

Piroth, Hans Georg & Janker, Peter M. (2004). Speaker-dependent differences in voicing and devoicing of German obstruents. *Journal of Phonetics*, 32, 81-109.

Pisoni, David B. (1997). Some thoughts on 'normalization' in speech perception. In K. Johnson & J. W. Mullennix (Hgg.), *Talker Variability in Speech Processing*. San Diego: Academic Press, pp. 9-32.

Pitt, Mark A. & McQueen, James M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language*, 39, 347-370.

Port, Robert F. & Crawford, Penny (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, 17, 257-282.

Port, Robert F. & Leary, Adam P. (2005). Against formal phonology. *Language*, 85, 927-964.

Port, Robert F. & O'Dell, Michael (1984). Neutralization of syllable final voicing in German. *Research in Phonetics: Report No. 4*, 93-134.

Port, Robert F. & O'Dell, Michael (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13, 455-471.

Port, Robert F. & Mitleb, Faris M. & O'Dell, Michael (1984). Neutralization of obstruent voicing in German is incomplete. In A. Koutsoudas (Hg.), *The*

Application of and Ordering of Grammatical Rules. The Hague: Mouton, pp. 55-73.

Manaster Ramer, Alexis (1996). A letter from an incompletely neutral phonologist. *Journal of Phonetics*, 24, 477-489.

Röttger, Timo & Winter, Bodo & Grawunder, Sven (2011). The robustness of incomplete neutralization in German. *Proceedings of the 17th ICPhS*, 1722-1725.

Röttger, Timo & Winter, Bodo & Grawunder, Sven (under review). Assessing the robustness of incomplete final voicing neutralization in German.

Röttger, Timo & Winter, Bodo & Kirby, James & Grice, Martine (2012). Incomplete Neutralization – a cognitive artifact of lexical co-activation? Poster presented at *Laboratory Phonology* 13, Stuttgart.

Sears, Chris R. & Hino, Yasushi & Lupker, Stephen J. (1995). Neighborhood size and neighborhood frequency-effects in word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 876-900.

Seidenberg, Mark S. & Tanenhaus, Michael K. (1979). Orthographic effects on rhyme monitoring. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 546-554.

Sereno, Joan A. & Jongman, Allard (1997). Processing of English inflectional morphology. *Memory and Cognition*, 25, 425-437.

Simonet, Miquel & Rohena-Madrazo, Marcos & Paz, Mercedes (2008). Preliminary evidence for incomplete neutralization of coda-liquids in Puerto Rican Spanish. In L. Colantoni, & J. Steele (Hgg.), *Selected Proceedings of the 3rd Conference on Laboratory Approached to Spanish Phonology*. Somerville, MA: Cascadilla Press, pp. 72-86.

Slowiaczek, Louisa & Dinnsen, Daniel A. (1985). On the neutralizing status of Polish word final devoicing. *Journal of Phonetics*, 13, 325-341.

Spivey, Michael J. (2007). *The continuity of mind*. New York: Oxford University Press.

Staum Casasanto, Laura & Jasmin, Kyle & Casasanto, Daniel (2010). Virtually Accommodating: Speech rate accommodation to a virtual interlocutor. In S. Ohlsson & R. Catrambone (Hgg.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*. Austin: Cognitive Science Society, pp. 127-132.

Taylor, Dennis Q. (1975). The Inadequacy of Bipolarity and Distinctive Features: The German "Voiced/Voiceless" Consonants. *LACUS forum*, 2, 107-119.

Trubetzkoy, Nikolai S. (1939). *Grundzüge der Phonologie*. Göttingen: Vandenhoeck and Ruprecht.

Trudgill, Peter (1974). *The social differentiation of English in Norwich*. Cambridge: Cambridge University Press.

van Oostendorp, Marc (2008). Incomplete devoicing in formal phonology. *Lingua*, 118, 1362-1374.

Warner, Natasha & Good, Erin & Jongman, Allard & Sereno, Joan (2006). Orthographic vs. morphological incomplete neutralization effects. *Journal of Phonetics*, 34, 285-293.

Warner, Natasha & Jongman, Allard & Sereno, Joan & Kemps, Rachèl (2004). Incomplete neutralization and other sub-phonemic durational differences in

production and perception: evidence from Dutch. *Journal of Phonetics*, 32, 251-276.

Wiese, Richard (1996). *The Phonology of German*. Oxford: Clarendon Press.

Wright, Richard (2004). Factors of lexical competition in vowel articulation. In J. J. Local & R. Ogden & R. Temple (Hgg.), *Laboratory Phonology* VI. Cambridge: Cambridge University Press, pp. 26-50.

Xu, Yi (2010). In defense of lab speech. *Journal of Phonetics*, 38, 329-336.

Yu, Alan C. L. (2007). Understanding near mergers: the case of mophological tone in Cantonese. *Phonology*, 24, 187-214.

Ziegler, Johannes C. & Ferrand, Ludovic (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin and Review*, 5, 683-689.

Zifonun, Gisela & Hoffmann, Ludger & Strecker, Bruno & Ballweg, Joachim & Braußе, Ursula & Breindl, Eva & Engel, Ulrich & Frosch, Helmut & Hoberg, Ursula & Vorderwülbecke, Klaus (1997). *Grammatik der deutschen Sprache*. Berlin: Walter de Gruyter.