



Letter to the Editor

Orthographic vs. morphological incomplete
neutralization effects

Natasha Warner^{a,b,*}, Erin Good^a, Allard Jongman^c, Joan Sereno^c

^a*Department of Linguistics, University of Arizona, P.O. Box 210028, Tucson, AZ 85721-0028, USA*

^b*Max Planck Institute for Psycholinguistics, Postbus 310, NL-6500 AH Nijmegen, the Netherlands*

^c*Linguistics Department, University of Kansas, Blake Hall, Lawrence, KS 66045, USA*

Received 16 June 2003; received in revised form 25 August 2004; accepted 3 November 2004

Abstract

This study, following up on work on Dutch by Warner, Jongman, Sereno, and Kems (2004. *Journal of Phonetics*, 32, 251–276), investigates the influence of orthographic distinctions and underlying morphological distinctions on the small sub-phonemic durational differences that have been called incomplete neutralization. One part of the previous work indicated that an orthographic geminate/singleton distinction could cause speakers to produce an incomplete neutralization effect. However, one interpretation of the materials in that experiment is that they contain an underlying difference in the phoneme string at the level of concatenation of morphemes, rather than just an orthographic difference. Thus, the previous effect might simply be another example of incomplete neutralization of a phonemic distinction. The current experiment, also on Dutch, uses word pairs which have the same underlying morphological contrast, but do not differ in orthography. These new materials show no incomplete neutralization, and thus support the hypothesis that orthography, but not underlying morphological differences, can cause incomplete neutralization effects.

© 2005 Elsevier Ltd. All rights reserved.

*Corresponding author. Department of Linguistics, University of Arizona, P.O. Box 210028, Tucson, AZ 85721-0028, USA. Tel.: +1 520 626 5591; fax: +1 520 626 9014.

E-mail address: nwarner@u.arizona.edu (N. Warner).

1. Introduction

This letter investigates the contribution of orthographic vs. underlying morphological differences to incomplete neutralization effects. It follows up on work by Warner, Jongman, Sereno, and Kems (2004), and provides evidence from word pairs with an underlying segmental difference but no orthographic difference. Furthermore, the underlying distinction investigated here exhibits global neutralization within the language, whereas most previous studies of incomplete neutralization have investigated positional neutralization phenomena.

1.1. Background on incomplete neutralization

Many researchers have determined that in several languages, when an underlying phonemic difference appears to be neutralized in a particular environment, speakers may still produce very small acoustic differences in the same direction as would be expected if the distinction were maintained (e.g., Charles-Luce, 1985; Dinnsen & Charles-Luce, 1984; Ernestus & Baayen, in press; Port & O'Dell, 1985; Slowiaczek & Dinnsen, 1985; and others cited in Warner et al., 2004). The best known example of this phenomenon, known as incomplete neutralization, is the apparent neutralization of final voicing in languages such as German, Polish, Catalan, and Dutch. German *Rat* 'advice' and *Rad* 'wheel' would both be transcribed as [ʁat], but researchers have found slightly longer vowel duration, shorter final burst duration, and longer closure voicing in *Rad* than in *Rat* (Port & O'Dell, 1985), differences which are in the same direction as the larger differences between intervocalic /t/ and /d/, where the voicing distinction is maintained. Often, such effects are not only small, but inconsistent across speakers or various conditions within an experiment, sometimes in unpredictable ways. However, a variety of work indicates that such effects do indeed exist.

Some researchers have suggested that incomplete neutralization effects stem from the influence of orthography, or of very careful speech. That is, speakers may attempt to make a distinction in a careful speech situation simply because the spelling suggests a difference. It is very difficult to remove the influence of orthography, particularly for well-educated speakers. However, a variety of experiments have found larger and more reliable incomplete neutralization effects in languages that mark the distinction orthographically, in tasks that involve reading, in more careful speech situations, or in situations where semantic information does not disambiguate pairs (Charles-Luce, 1993; Fourakis & Iverson, 1984; Jassem & Richter, 1989; Kopkalli, 1993; Port & Crawford, 1989). Smaller effects, or no difference at all, are found in less careful speech situations, in tasks which do not emphasize orthography, or in languages where the underlying difference is not reflected in the orthography.

1.2. Previous findings of Warner et al. (2004)

Warner et al. (2004) investigated final devoicing in Dutch, for which no large studies on this topic were previously available. They found that in words such as *meed* 'avoided' and *meet* 'measures' (a neutralization environment), vowel duration is significantly longer before underlying /d/ than /t/, both for words with phonemically long vowels (such as *meed/meet*) and words with phonemically short vowels (such as *rat* 'rat' and *rad* 'wheel'). The difference was significant both

by subjects and by items, but the average difference was only 3.5 ms. Burst duration was significantly longer for underlying /t/ words than /d/ words, but only in the words with phonemically long vowels. (This type of inconsistency across conditions is common in studies of incomplete neutralization.) This basic result extends the incomplete neutralization of final devoicing effect to Dutch (as also found by Ernestus and Baayen, *in press*).

Warner et al. (2004) performed a separate experiment to investigate the influence of orthography in incomplete neutralization. This experiment used pairs such as *heten* ‘to be called’ and *heetten* ‘were called,’ which differ systematically in their spelling, but both represent the phoneme string /hɛ:tən/. Thus, this experiment examined an apparent geminate/singleton neutralization, rather than the more commonly investigated final voicing neutralization. Consonant duration was significantly longer for words spelled with a double consonant (significant both by subjects and items, with a 3.4 ms average difference). Vowel duration had a significant effect only for words with medial /t/ (e.g., *heten/heetten*), not for words with medial /d/ (e.g., *kleden* ‘to dress’/ *kleedden* ‘they dressed’). For the medial /t/ words, vowels were significantly shorter, by 3.0 ms, in words spelled with VVtt than in Vt words. Warner et al. (2004) suggested that these effects were caused by the orthographic difference, and that the shorter vowel duration in VVtt words reflected the syllable structure of the spelled words.

1.3. Purpose of the current study

Although the purpose of the experiment above using pairs such as *heten/heetten* was to test a situation with an orthographic but no underlying phonemic difference, there is an alternate possible interpretation of those materials. The reason for the orthographic difference between such words is that the plural past tense form (e.g., *heetten*) contains the past tense allomorph /-t/.¹ The underlying string of morphemes in *heetten* is /hɛ:t-t-ən/, while *heten* is simply /hɛ:t-ən/. The phonemic representations of the two words are thus different at some level, although one could also consider the past tense morpheme to have a null allomorph after the phoneme /t/, leaving both with the same underlying form. It is possible, therefore, that the durational differences found in the previous experiment do not reflect orthographic influence but are instead another example of incomplete neutralization of an underlying phonemic distinction. This case would involve neutralization of geminate /tt/ (or /dd/) with singleton /t/ (or /d/). This would be a somewhat different case from neutralization of final voicing: voicing is distinguished in other environments (intervocally and in onset), whereas Dutch does not have geminates within the prosodic word in any environment (Booij, 1995). Thus, neutralization of a geminate/singleton distinction is a global neutralization, and is not the same as the positional neutralization phenomena that have shown incomplete neutralization effects before.

We do not know of any pairs of words in Dutch which differ systematically in spelling and yet provide a more convincing case of having the same underlying phoneme string.² Therefore, we designed the current experiment to determine whether pairs of words which differ in their

¹Because of the addition of the second orthographic consonant, the preceding vowel has to be doubled in order to indicate that it is a phonemically long vowel despite the fact that it is followed by two consonants (Nunn, 1998).

²There are some pairs which differ in orthography but not phonemic string, such as *mail* ‘e-mail message’ vs. *meel* ‘flour’ and *ijs* ‘ice’ vs. *eis* ‘demand.’ Some such pairs involve recent borrowings from other languages, while others depend on the fact that a few phonemes of Dutch can be spelled more than one way (*ei/ij*, both of which represent /ei/

underlying morphemic string, but not in orthography, would show similar incomplete neutralization effects. For this experiment, we chose words such as *heet* ‘be called’ following the pronouns *ik* ‘I’ and *hij* ‘he.’ In *ik heet* ‘I am called,’ *heet* consists of /hɛt/ with no suffix. In *hij heet* ‘he is called,’ *heet* consists of /hɛ:t-t/, with an additional third person singular morpheme /-t/. Thus, the difference in underlying phoneme strings, and the phonological status of the neutralization, are extremely similar to the pair *heten* /hɛ:t-ən/ vs. *heetten* /hɛ:t-t-ən/, but in this case, there is no difference in the orthography.

Our prediction is that speakers will produce complete neutralization (no significant duration differences) in pairs with an underlying morphological difference but no orthographic difference (e.g., *ik heet/hij heet*). This would support the hypothesis that orthographic differences can cause effects comparable to incomplete neutralization effects, and would support the idea that a highly abstract underlying difference (involving global neutralization) cannot cause such effects. Alternatively, if speakers produce small durational differences even in these materials, particularly if those differences include longer consonant duration for the underlying double consonant condition (*hij heet*, with additional /-t/ morpheme), this would indicate that the effect in the previous study was caused by the underlying rather than the orthographic difference. It would also suggest that even a rather abstract global neutralization phenomenon, such as degemination in a language with no geminates, can cause incomplete neutralization.

2. Methods

2.1. The current study

For this experiment, we chose 20 verbs of Dutch with /t/ as the final segment of the stem. Ten had phonemically long vowels (e.g., *heet*) and 10 had phonemically short vowels (e.g., *zit* ‘sit’). Only verbs with /t/ as the final stem segment can be used, as these are the only verbs which have the same spelling with and without the third person singular suffix /-t/. The materials appear in Table 1. Each verb was placed once after *ik* ‘I’ and once after *hij* ‘he.’ The resulting 40 items were arranged in a pseudo-random order (the same word never occurring twice in a row) into 4 lists of 10 items each, with an additional filler item at the beginning and end of each list. The pronouns were written in a slightly smaller font than the target words.

Fifteen native Dutch speakers who had not participated in any of the experiments reported by Warner et al. (2004) served as speakers.³ Recordings were made in the Netherlands, and the

(footnote continued)

in the standard dialect). These pairs do not represent a systematic difference, and it is difficult to make a clear prediction for how they might influence duration.

³If one performed both the *heten/heetten* experiment and the *ik heet/hij heet* experiments with the same speakers, one could arguably be even more certain of any difference in results for the two experiments. Thus, it might seem desirable to use the same speakers rather than entirely new ones. However, having just read a list of words such as *heten* and *heetten* could influence the speakers’ productions of the subsequent list of words such as (*ik*) *heet* and (*hij*) *heet*, or vice versa. Throughout this series of experiments, we have attempted to avoid any potential influence of having done one experiment on the results of the next. Furthermore, the two production experiments in the Warner et al. (2004) study also used different speakers, but still showed nearly identical patterns and sizes of results. Thus, any difference in results between the current study and the previous ones is likely to be reliable, even with different speakers.

Table 1
Materials

Phonemically short vowels		Phonemically long vowels	
ik/hij pit [pɪt]	I/he sleep(s)	ik/hij weet [we:t]	I/he know(s)
ik/hij knot [knɔt]	I/he knot(s)	ik/hij laat [la:t]	I/he let(s)
ik/hij zit [zɪt]	I/he sit(s)	ik/hij schiet [sxɪt]	I/he shoot(s)
ik/hij zet [zɛt]	I/he set(s)	ik/hij blaas [bla:t]	I/he bleat(s)
ik/hij schat [sxat]	I/he estimate(s)	ik/hij giet [xɪt]	I/he pour(s)
ik/hij dut [dʏt]	I/he nap(s)	ik/hij eet [e:t]	I/he eat(s)
ik/hij jat [jɑt]	I/he steal(s)	ik/hij heet [he:t]	I/he am/is called
ik/hij plet [plɛt]	I/he flatten(s)	ik/hij loot [lo:t]	I/he draw(s) lots
ik/hij spat [spɑt]	I/he spatter(s)	ik/hij haat [ha:t]	I/he hate(s)
ik/hij spit [spɪt]	I/he dig(s)	ik/hij boet [but]	I/he repent(s)

recording session was carried out entirely in Dutch by a native Dutch-speaking assistant. Speakers were instructed to read the entire two-word phrase (e.g., *ik heet*) to themselves silently, but to read aloud only the verb (e.g., *heet*). Since the effects found in the previous experiments were extremely small (significant effects of approximately 3.5 ms), if speakers read the entire two-word phrase aloud, it is likely that the preceding /k/ in *ik* and preceding diphthong in *hij* would create larger differences in duration of the following vowel than any potential incomplete neutralization effect. (In the previous study, control factors such as phonemic vowel length often produced significant but unpredicted durational differences larger than the incomplete neutralization effects. In unpublished pilot work, we also found that any differences in surrounding environment could produce larger effects than incomplete neutralization, despite all attempts at segmental and prosodic control.)⁴ Each speaker read all the lists twice, once in each of two different orders. The speech was recorded on a DAT recorder using high quality equipment, in a sound protected booth, and was digitized at 16,000 Hz and analyzed using the ESPS/XWaves software.

For each item, vowel duration, final stop closure duration, and total duration of the final consonant were measured. The end of the vowel was defined as the end of F2, and closure duration was defined as the time from end of the vowel to beginning of the burst. End of the burst was defined as the end of visible noise louder than background noise in the waveform. The onset of the vowel was defined as onset of voicing after voiceless obstruents and onset of F2 after voiced

⁴The method of having subjects read a two-word phrase silently for context, but read only the target item out loud, is unusual, and it could potentially prevent speakers from realizing an incomplete neutralization effect if one is present. However, we feel this method is necessary because of the danger of differences in preceding environment affecting duration. Furthermore, this method makes the two experiments more similar in recording method, by having the target items read in isolation in both. We are not aware of this method having been used before: in the previous work on Catalan (Charles-Luce 1993; Dinnsen & Charles-Luce 1984), which does not mark the underlying difference orthographically, speakers read entire paragraphs containing the target items, which leaves open the danger of effects of surrounding environment or prosody. If one could find a systematic orthographic distinction with no underlying distinction whatsoever in some other language, or a way to make sure that speakers adequately activate the correct form in a case such as *ik heet/hij heet* without producing the environment, either of these would provide an excellent way to confirm the current findings.

obstruents. The boundary between nasals or /l/ and vowels was determined by the sudden change in energy distribution in the spectrogram. Onset of the vowel after glides, for the few items beginning with glides, was defined as a sudden increase in amplitude if present, or if not, as half way through the F2 transition. For the previous experiment on doubled consonants (the *heten/heetten* experiment), only vowel duration and total consonant duration were measured. Since the current experiment requires that the final consonants be utterance-final, and this makes it more difficult to measure total consonant duration, closure duration, which can be measured far more accurately, was also measured. All measurements were made by the second author in order to maintain consistency.

2.2. Interpretation of a null effect

Our prediction is for complete neutralization, thus, for a null effect. It is of course very difficult to demonstrate that a null effect is meaningful. However, in this case, two closely related previous experiments have shown that very small effects (approximately 3.5 ms difference) reach statistical significance. Therefore, we modeled the experimental design and methods for the current experiment very closely on those previous studies: the same number of speakers and items were used, and the second factor (phonemic vowel length) is the same as in one of the previous experiments. (The other, the *heten/heetten* experiment, had voicing of the stem-final consonant instead of vowel length as the second factor, because that orthographic alternation exists only for verbs with phonemically long stem vowels.) The speakers were recorded under the same circumstances, using the same word list-reading method. This extremely careful speech situation may make speakers more likely to produce sub-phonemic differences. (In both experiments, speakers read the target items as single words in isolation. In the current experiment, the preceding pronoun *ik/hij* had to be written in the materials to indicate which word was intended, but the actual production of target words in isolation is identical to the previous study.)

Much of the earlier literature on incomplete neutralization used only a few speakers and a smaller number of items, and found inconsistent results. In Warner et al. (2004), we proposed that such small effects as incomplete neutralization require a large number of speakers and items in order to detect them (or rule them out) reliably. Because our previous studies were able to detect such effects reliably, if the current experiment, using the same large number of speakers and items, shows a null effect, it can indeed be assumed that the current manipulation does not produce an incomplete neutralization effect, at least not one comparable in size to the 3.5 ms differences found in the previous work.

3. Results

Data was averaged across the two productions of each item by each speaker, and was then averaged across items for the by-subjects analysis and across subjects for the by-items analysis. Each of the dependent variables vowel duration, total consonant duration, and consonant closure duration were analyzed using by-subjects and by-items two-factor repeated measures ANOVAs, with morphophonological form (final /t-t/ or final /t/) and phonemic vowel length as the two factors. For all three dependent variables, the main effect of phonemic vowel length was

Table 2

Average durations for each dependent variable and condition, in milliseconds

Dependent var.	Phonemically short vowel		Phonemically long vowel	
	Underlying /t/ (<i>ik zit</i>)	Underlying /t-t/ (<i>hij zit</i>)	Underlying /t/ (<i>ik heet</i>)	Underlying /t-t/ (<i>hij heet</i>)
Vowel duration	108.4	108.5	165.9	167.0
Closure duration	96.9	96.2	89.0	89.5
Consonant duration	204.1	206.1	198.1	197.1

An example of each condition appears under the condition's heading.

Table 3

Summary of statistical results

Dependent variable	Effect	By subjects test (df for all tests = 1,14)	By items test (df for all tests = 1,9)
Vowel duration	Length	$F = 348.58, p < .001$	$F = 11.15, p < .01$
	Morph.	$F = .48, p = .502$	$F = .65, p = .439$
	Length \times Morph.	$F = .24, p = .633$	$F = .47, p = .510$
Closure duration	Length	$F = 28.70, p < .001$	$F = 24.57, p < .005$
	Morph.	$F = .03, p = .857$	$F = .01, p = .933$
	Length \times Morph.	$F = .86, p = .369$	$F = .47, p = .508$
Consonant duration	Length	$F = 13.13, p < .005$	$F = 21.29, p < .005$
	Morph.	$F = .21, p = .656$	$F = .08, p = .788$
	Length \times Morph.	$F = .95, p = .346$	$F = 1.08, p = .326$

Main effects of phonemic vowel length (“Length”) and morphophonological form (“Morph.”), as well as their interaction, are shown. Exact F ratio and p levels are given even for F ratios less than 1, since the lack of an effect is at issue.

significant both by subjects and by items, but neither the main effect of morphophonological form nor the interaction of the two factors was significant either by subjects or by items. Words with phonemically long vowels have longer vowel duration and shorter consonant duration and closure duration than those with phonemically short vowels. Table 2 shows the results for each dependent variable. Table 3 summarizes the statistical tests.

Because this is a null effect, it is important to consider the power of the experiment to detect any potential effect. It is possible to make a rather accurate estimate of the power the current experiment would have to detect an effect of comparable size to those in the previous experiments (approximately 3.5 ms difference between groups), since we have done two previous experiments with the same number of subjects and items, the same recording procedure, and similar materials. The basic incomplete neutralization experiment discussed above (using words such as *meed/meet*) had power of .81 to detect the vowel duration difference (3.5 ms) in that experiment. The previous orthographic difference experiment (using words such as *heten/heetten*) had power of .88 to detect

the consonant duration difference (3.4 ms). The current experiment should have a similar power to detect effects of comparable size. (The error terms for these effects are comparably small for all three experiments.) Furthermore, the current experiment has power of .99 to detect the 7.3 ms effect of phonemic vowel length on consonant closure duration found in this experiment, and power of .92 to detect the 7.5 ms effect of phonemic vowel length on total consonant duration. The previous experiments' high power to detect effects of approximately 3.5 ms and the current experiment's high power to detect larger but still small differences both suggest that the lack of effect of underlying morphophonological form in the current experiment is reliable.

4. Discussion

The current experiment suggests that when the only difference between two forms is a difference in the underlying morpheme strings, such as in /hɛɪt/ (1st person) vs. /hɛɪt-t/ (3rd person), there is complete neutralization. Speakers do not seem to produce any reliable durational differences in such pairs, even of the very small size found in previous experiments. This leads us to speculate that in the previous experiment with words such as *heten* /hɛɪt-ən/ vs. *heetten* /hɛɪt-t-ən/, the reliable durational differences we found were caused by the orthographic difference and not the underlying morphological difference. In combination, these studies show that orthographic differences can cause effects that strongly resemble incomplete neutralization, but differences in the underlying morpheme string apparently cannot.

This has two implications for the phenomenon of incomplete neutralization in general. First, the current findings strengthen the evidence that there is an orthographic component to incomplete neutralization. As discussed above, various studies show a general tendency for smaller incomplete neutralization effects, or for complete neutralization, in tasks which de-emphasize orthography and in languages which do not represent the underlying difference orthographically. The current results are compatible with an account in which an underlying difference with no accompanying orthographic difference does not lead to incomplete neutralization, while an identical underlying difference which is orthographically encoded does.

Second, the underlying difference tested here, an underlying geminate/singleton difference created by concatenation of morphemes in a language which has no geminates, is of a rather different type than the final voicing neutralization that has been most widely tested for incomplete neutralization. Final devoicing is positional neutralization, and the distinction is maintained in other environments in the same language. Thus, speakers are familiar with the voicing distinction and with both the voiced and voiceless phonemes. However, since Dutch has no surface geminates within the prosodic word (Booij, 1995), Dutch speakers are not familiar with a double vs. single consonant distinction. (Booij (1995) states that any sequence of identical consonants must degeminate within the prosodic word, and optionally degeminates at higher levels.) It is not entirely clear that speakers produce incomplete neutralization in the absence of an orthographic difference even for final devoicing (cf. conflicting results on Catalan and Turkish, neither of which spell the final voicing distinction (Charles-Luce, 1993; Dinnsen & Charles-Luce, 1984; Kopkalli, 1993)). Incomplete neutralization may be entirely caused by orthographic differences, but if it is not, the current experiment suggests that it may be restricted to positional neutralization phenomena, whereas global neutralization leads to complete neutralization.

Acknowledgments

This research was supported by a SPINOZA award from the Nederlandse Organisatie voor Wetenschappelijk Onderzoek to Anne Cutler. We are very grateful to Mirjam Ernestus, Mike Hammond, Anne Cutler, and Tom Bever for discussion of this material, and to two anonymous reviewers for their helpful comments. We would also like to thank Evelien Akker and Keren Shatzman for their help in collecting and analyzing parts of the data, and Chris Carpenter for technical support. Any errors or misinterpretations are, of course, our own.

References

- Booij, G. (1995). *The phonology of Dutch*. Oxford: Oxford University Press.
- Charles-Luce, J. (1985). Word-final devoicing in German: Effects of phonetic and sentential contexts. *Journal of Phonetics*, 13, 309–324.
- Charles-Luce, J. (1993). The effects of semantic context on voicing neutralization. *Phonetica*, 50, 28–43.
- Dinnsen, D., & Charles-Luce, J. (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics*, 12, 49–60.
- Ernestus, M., & Baayen, H. (in press). *The functionality of incomplete neutralization in Dutch: The case of past tense formation*. Papers in Laboratory Phonology VIII.
- Fourakis, M., & Iverson, G. (1984). On the ‘incomplete neutralization’ of German final obstruents. *Phonetica*, 41, 140–149.
- Jassem, W., & Richter, L. (1989). Neutralization of voicing in Polish obstruents. *Journal of Phonetics*, 17, 317–325.
- Kopkalli, H. (1993). *A phonetic and phonological analysis of final devoicing in Turkish*. Unpublished Dissertation, The University of Michigan.
- Nunn, A. M. (1998). *Dutch orthography: A systematic investigation of the spelling of Dutch words*. Ph.D. dissertation. The Hague: Holland Academic Graphics.
- Port, R., & Crawford, P. (1989). Incomplete neutralization and pragmatics in German. *Journal of Phonetics*, 17, 257–282.
- Port, R., & O’Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 13, 455–471.
- Slowiaczek, L., & Dinnsen, D. (1985). On the neutralizing status of Polish word—final devoicing. *Journal of Phonetics*, 13, 325–341.
- Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, 32, 251–276.