

The Recognition of Phonologically Assimilated Words Does Not Depend on Specific Language Experience

Holger Mitterer^{a,b}, Valéria Csépe^c, Ferenc Honbolygo^c, Leo Blomert^a

^a*Unversiteit Maastricht, Faculty of Psychology, Department of Cognitive Neuroscience*

^b*Max Planck Institute for Psycholinguistics, Nijmegen*

^c*Hungarian Academy of Sciences, Department of Psychophysiology,
Research Group of Developmental Psychophysiology*

Received 9 December 2004; received in revised form 8 September 2005; accepted 12 September 2005

Abstract

In a series of 5 experiments, we investigated whether the processing of phonologically assimilated utterances is influenced by language learning. Previous experiments had shown that phonological assimilations, such as /lean#bacon/ → [leam bacon], are compensated for in perception. In this article, we investigated whether compensation for assimilation can occur without experience with an assimilation rule using automatic event-related potentials. Our first experiment indicated that Dutch listeners compensate for a Hungarian assimilation rule. Two subsequent experiments, however, failed to show compensation for assimilation by both Dutch and Hungarian listeners. Two additional experiments showed that this was due to the acoustic properties of the assimilated utterance, confirming earlier reports that phonetic detail is important in compensation for assimilation. Our data indicate that compensation for assimilation can occur without experience with an assimilation rule, in line with phonetic–phonological theories that assume that speech production is influenced by speech-perception abilities.

Keywords: Speech perception; Phonological assimilation; Phonology; Continuous speech; Hungarian; Dutch

1. Introduction

How is a spoken word recognized? According to the predominant view, the acoustic input is first transformed into a more abstract code. This abstract code is then used to achieve lexical access (McQueen & Cutler, 2001). Although the grain size of the basic unit of this code is still under debate—with proposals ranging from phonological features (Marslen-Wilson & Warren, 1994) to demisyllables (Massaro, 1998)—most psycholinguistic approaches converge on the assumption of intermediate representations of a certain amount of abstraction in spo-

Correspondence should be addressed to Holger Mitterer, Max-Planck Institute for Psycholinguistics, Postbus 310, 6500 AH Nijmegen, The Netherlands. E-mail: holger.mitterer@mpi.nl

ken-word recognition. However, all of these proposals put a rather heavy load on the assumed preprocessing mechanisms that have to transform the acoustic input into some form of such an intermediate representation. Achieving an intermediate, phoneme-like representation of an auditory input is hampered by a number of nonlinguistic influences on the speech signal. The acoustic implementation of, for instance, a phoneme, depends among other influences, on the speech rate, the surrounding phonemes (Farnetani, 1997; Gussenhoven & Jacobs, 1998), the likelihood of the word in the sentence context (Jufursky, Bell, & Girand, 2002), as well as the height, sex, age, emotional, and sociological status of the speaker (Ladefoged, 1996; Lerner & Loewe, 1956; Traunmüller & Eriksson, 2000; van Bezooijen, 1995). Given this list of variation-adding influences, it becomes evident that the preprocessing mechanisms, which transform the acoustic input into an intermediate representation before lexical access, face a rather complicated task.

However, even if the acoustic realization of a phoneme varies strongly due to the variation-adding influences, an intermediate representation may still be achieved by a context-sensitive phoneme-detection mechanism. Indeed, for all influences listed previously, which add variance to the speech signal, evidence has been found for a context-sensitive interpretation of the speech signal (Beddor & Krakow, 1999; Evans & Iverson, 2004, Fowler & Brown, 2000; Fowler, Brown, & Mann, 2000; Gaskell & Marslen-Wilson, 1996, 1998; Gow, 2002, 2003; Ladefoged & Broadbent, 1957; Lotto & Kluender, 1998; Lotto, Kluender, & Holt, 1997; Mann, 1980; Mann & Repp, 1981; Nearey, 1989; Sussman, Fruchter, Hilbert, & Sirosh, 1998). These studies have probed context sensitivity of the speech-perception system by presenting a given stimulus in different acoustic and phonetic contexts. In all these studies, it has been shown that the speech-perception system is context sensitive so that the percept derived from a speech sound depends on its context. In addition, the context sensitivity is usually “compensatory.” The compensatory context effects in speech perception are the inverse of the variation-adding influences in speech production. For instance, as a vowel may become nasalized in production due to coarticulation with an adjacent nasal consonant, in perception, some nasalization is parsed from a vowel if it is accompanied by a nasal consonant (see, e.g., Beddor & Krakow, 1999). In effect, this may yield a more or less stable phoneme-like code that is resistant to context-dependent perturbations occurring during production.

Another instance of such an inverse relation between speech production and speech perception is observed for the case of phonological assimilations. Phonological assimilations are driven by context-dependent rules that regulate how a phoneme may be influenced by a following or preceding phoneme. Such assimilation rules are language specific (Gussenhoven & Jacobs, 1998), although some general tendencies were also reported (Kohler, 1990; Steriade, 2001). A typical exemplar of an assimilation rule is the rule for coronal place assimilations in English. According to this rule, word-final coronal stops and nasals may take over the place of articulation of the subsequent phoneme. Hence, the word *lean* may be pronounced as “*lean*” in “*leambacon*,” but not in “*leamsalami*.” In perception, the opposite effect has been observed. A word-final [m] is sometimes perceived as an /n/ in “*leambacon*” but less so in “*leamsalami*” (Gaskell & Marslen-Wilson, 1996, 1998; Gow, 2002, 2003; Mitterer & Blomert, 2003). Hence, the context-sensitive change in production—/n/ may become [m] before /b/, but not before /s/—is countered by an inverse context-sensitive change in perception—[m] may become /n/ before /b/ but not before /s/.

Evidence for the context-sensitive interpretation of assimilated utterances stems from a multiplicity of tasks from cross-modal priming (Gaskell & Marslen-Wilson, 1996, 2001; Gow, 2002, 2003) phoneme monitoring (Gaskell & Marslen-Wilson, 1998), and simple identification tasks (Mitterer & Blomert, 2003). Given ample evidence for a process of compensation for assimilation, the question arises how this compensation for assimilation mechanism works. A first hint to the nature of this mechanism may be derived from the way in which we acquired this mechanism; did we inherit it or does everybody have to learn how to deal with the specific assimilation rules in one's own native language during development? The latter account is more prominent in models of language acquisition (e.g., Pepperkamp, 2003). A recently proposed neural-network model for the perception of assimilation follows this learning hypothesis (Gaskell, 2003) and shows that a covariate learning mechanism may be sufficient to acquire compensation for assimilation. This assertion gains creditability by the fact that learning has also been shown to be important for other context effects in speech perception (e.g., Holt, Lotto, & Kluender, 2001; Howell & Rosen, 1984; Kuhl, 1992; Sinnott & Saporita, 2001).

However, there is also evidence that not only learning but also basic auditory processes are instrumental in many speech-related phenomena, for instance, compensation for coarticulation (e.g., Cutting & Rosner, 1974; Delgutte, 1997; Lotto & Kluender, 1998; Lotto et al., 1997; Pastore 1987). It is therefore necessary to explore the possibility that auditory processes may also be instrumental in compensation for assimilation, and accordingly, compensation for assimilation may be independent of specific language experience. Such a view resonates with a theoretical perspective developed within phonetics–phonology: Jakobson, Fant, and Halle (1963, p. 13) argued that “we speak to be heard to be understood.” One kind of evidence that we speak to be heard comes from the design of vowel systems in different languages. Although languages may differ vastly in the number of vowels they use, vowel systems with the same number of vowels tend to be similar. Moreover, this similarity seems to be explained by a preference for peripheral vowels in the vowel space so that pairs of vowels are maximally distinguishable (e.g., Schwartz, Boë, Vallée, & Abry, 1997). Optimizing speech for maximal perceptual distinctiveness obviously does not explain articulatory simplifications such as phonological assimilations. Lindblom (1990, 1996) argued that not only “ease of perception” but also “ease of production” determines phonological systems. Hence, assimilation may be explained by “ease of articulation” because *assimilation* by definition means that the sequence of assimilated and assimilating segments will share some phonological feature and will therefore be easier to produce. These articulatory simplifications, in turn, seem to be constrained by perception. Kohler (1990) noted that place assimilations occur more often in the world's languages for stops and nasals than for fricatives. He argued that this indicates a perceptual constraint on assimilation rules, because place of articulation is supposedly more easy to perceptually distinguish in fricatives than in nasals and stops. Hura, Lindblom, and Diehl (1992) backed up this assertion with experimental results: Listeners find it more difficult to distinguish place of articulation in final nasals and stops than in fricatives. Accordingly, the most often attested nasal place assimilation blurs a phonemic distinction that listeners find difficult to exploit reliably in the first place. This has led to a view that phonological assimilations tend to be shaped by language-independent perceptual preferences (see, e.g., Hume & Johnson, 2001; Hura et al., 1992; Kohler, 1990; Ohala, 1990; Steriade, 2001): Assimilations are perceptually inconspicuous articulatory simplifications. If this is the case, it follows that there is little need

for the listeners to adapt to assimilation rules, because assimilation rules are adapted to the listeners. Thus, similarly as for coarticulation, compensation for assimilation may be rooted in auditory processes independent of specific language experience. The purpose of this study is to examine whether listeners eventually can compensate for assimilation without experience with a given assimilation.

Some studies do not support this seemingly straightforward phonetic–phonological account. Cross-linguistic investigations on the perception of phonological assimilation found evidence for language-specific perception. Otake, Yoneyama, Cutler, and Van der Lugt (1996) showed that Japanese, but not Dutch, listeners were able to use nasal place assimilation in Japanese words to predict the postassimilation context. Weber (2001) showed that German, but not Dutch, listeners responded with a pop-out effect to violation of a German fricative assimilation rule. However, both situations investigate cases in which compensation for assimilation is not strictly necessary, because the assimilation rule does not add to the invariance problem in these cases. That is so because the assimilation rules investigated by Otake et al. and Weber do not even superficially blur a phonemic distinction in the languages examined. For the case of the German fricative assimilation, the assimilation rule applies obligatorily to back fricatives, which surface as velar [x] after back vowels and as palatal [ç] after front vowels. German, however, does not distinguish velar and palatal fricatives as different phonemes. This is similar for the case of Japanese nasal place assimilation. Japanese does not distinguish nasals according to place of articulation, and assimilated forms are therefore not possible different word forms affording compensation. However, place assimilation in English makes an /n/ quite similar to an /m/, which is a different phoneme in English. Another well-cited study by Lahiri and Marslen-Wilson (1991) investigated assimilatory vowel nasalization cross-linguistically. They found that English, but not Bengali, listeners use vowel nasalization to predict a following nasal consonant. In Bengali, but not in English, vowel nasalization is phonemic. Therefore, this study also does not necessarily reflect a listener's adaptation to native-language assimilation rules, but may rather reflect, as do the studies by Weber and Otake et al., an effect of the different phonemic inventories of the two languages.

Two other studies are more relevant for this case. Darcy (2002) presented voice-assimilated French tokens to French and English listeners in a sentence. The task was to spot a prespecified target word in the sentence. Only French and English listeners with a very good command of the French language were able to spot the words in their assimilated form in an appropriate context. This seems to indicate that the context-sensitive perception of assimilated forms needs to be learned.

More in line with the language-independent account, Gow and Im (2004) did not find evidence for a language-specific processing of assimilated utterances. They presented assimilated Hungarian and Korean utterances to English as well as Hungarian and Korean listeners. They investigated the perception of the postassimilation context. According to the feature-parsing account of Gow (2003), the evidence for the conflicting place or voice information in the assimilated segment is parsed from that segment and attributed to the assimilating context, that is, the postassimilation context in the case of the regressive assimilation rules. Feature parsing has two consequences. First, the assimilated segment is “freed” from disturbing feature cues indicating a “wrong” feature value. Second, the postassimilation context can be recognized faster, because evidence for its presence is already picked up from the assimilated segment.

Accordingly, faster monitoring for postassimilation context may be seen as evidence that compensation for assimilation occurred. Gow (2003) indeed showed that regressive and progressive context effects go hand in hand for coronal place assimilation in English: Faster monitoring for postassimilation context was observed for the same token, for which regressive context effects occurred. Gow and Im went on to show that monitoring for postassimilation context for Hungarian voicing assimilation (voiceless fricative becomes voiced in front of voiced stops) and Korean labial-to-velar place (labials assimilated to velars) assimilation was similar for native speakers of Hungarian and Korean on the one hand and native speakers of English on the other hand. For the case of the Hungarian voicing assimilation, both Hungarian and English listeners were faster to monitor for a /d/ after a partly assimilated fricative, in line with the assumption of the feature-parsing model that the partial voicing of the fricative is used as a cue for voicing in the following stop. For the case of the Korean labial-to-velar assimilation, monitoring for the postassimilation context was not influenced by an assimilated preceding context. Gow and Im argued convincingly that the difference in results for the two different assimilation phenomena may be related to the acoustic properties of the assimilations involved, which indicate that Hungarian voicing assimilation may be a gradient phenomenon—leading to partially voiced fricatives—whereas Korean labial-to-velar assimilation may be categorical. Nevertheless, the evidence that English listeners compensate for assimilation rules that do not occur in English is not overwhelming. For the case of Hungarian voicing assimilation, the evidence points in a positive direction. However, the data of Gow and Im addressed the question of the perception of the assimilated segment only indirectly, by assessing the perception of the assimilating segment. For the Korean rule, no evidence for facilitated monitoring for postassimilation context after assimilated utterances was observed, which seems to prompt the unlikely conclusion that Korean listeners cannot deal with Korean labial-to-velar assimilation. In summary, the data of Gow and Im showed that the perception of assimilated utterances may be independent of native-language experience. At the same time, their evidence addressed the question of compensation-for-assimilation by non-native speakers only indirectly, by assessing the perception of the assimilating segment and not the perception of the assimilated segment. The question of whether compensation for assimilation can occur without experience with a given assimilation rule thus remains unsettled.

These experiments investigate cross-linguistically whether listeners eventually can compensate for assimilation without experience with a given assimilation. Such a finding would be in line with the assumption that assimilation rules are adapted to the listener. To study the perception of assimilations by naive listeners, it is necessary to ensure that these listeners have not been exposed to the kind of assimilation under study. This constraint cautions against the use of tokens that undergo nasal place assimilation (used in the Mitterer & Blomert, 2003, study), because this rule not only exists in Dutch but also in, among many other languages, English. Therefore, we used a, in Dutch, nonexistent assimilation from the Hungarian language. Hungarian belongs to the Finno-Ugric language family that contains major exponents—Finnish, Estonian, and Hungarian—rarely acquired by non-native learners. The assimilation rule used in this study was a rule of liquid assimilation (see Olsson, 1992, p. 57; Siptár & Törkenczy, 2000, p. 182): the Hungarian word for *left* [bɔl]—and all other words ending on /l/—may be pronounced with a word-final [r] (hence, [bɔr]) if the next phoneme is also an /r/ (as in [bɔlro:l], Engl. “from the left”). However, this change from an alveolar lateral /l/ to an alveolar

trill [r] is not allowed to happen in [bɔlna:l] (Engl. “at the left”). For naive listeners, we used Dutch listeners, who did not have any knowledge of Hungarian. Dutch listeners can be assumed to be naive with regard to liquid assimilation, because liquid assimilation does not occur in Dutch. Although a contrast between /l/ and /r/ does occur in Dutch, neither standard Dutch phonology (Booij, 1995) nor a corpus investigation of casual Dutch speech (Ernestus, 2000) reported an assimilation of /l/ to [r]. It should also be noted that the implementation of, especially, /r/ is subject to large geographic variations in the Netherlands, with the alveolar trill, which is used in Hungarian, largely on the retreat in postvocalic position. Therefore, it is very unlikely that Dutch listeners are familiar with any kind of assimilation such as the Hungarian liquid assimilation. In the first experiment, we tested whether Dutch listeners nevertheless compensate for Hungarian liquid assimilation the way they compensate for Dutch nasal place assimilation. In the next three experiments we validated the effects by testing Hungarian listeners with the same stimuli and investigated the significance of phonetic detail by using slightly different language materials.

One reason for the paucity of cross-linguistic evidence may lie in the nature of the tasks often used to test compensation for assimilation. Tasks such as cross-modal identity priming, word monitoring, and phoneme monitoring require the listener to attend carefully to the speech stream. Especially, this last requirement may render these paradigms not ideally suited for cross-linguistic comparisons, because non-native language material often sounds quite unintelligible at a first encounter (e.g., it is a fairly common experience that one has trouble isolating words, let alone phonemes, in streams of an unknown language). A possible solution to this problem is to apply native and non-native assimilations to the respective native language of the listeners in a cross-linguistic design. In this case, unfortunately, the cross-linguistic comparison is confounded with the phonetic detail in the stimuli. This is problematic, because Gow (2002) has shown that the exact phonetic implementation of the phonological assimilation influences compensation for assimilation.

Another possible solution comes from the study by Mitterer and Blomert (2003). They investigated this “compensation for assimilation” using mismatch negativity (MMN). MMN is a component of the event-related potentials (ERPs) that arise when listeners are confronted with a so-called oddball series, in which an often-presented stimulus—the standard—contrasts with a seldom-presented stimulus—the deviant. The listener is not required to pay attention to the auditory stimuli and is usually distracted by a silent movie or by reading a book. In such a design, the deviant can be observed to yield a more negative ERP at frontocentral electrodes 100 to 200 msec after the onset of the acoustic mismatch between standard and deviant, which is often accompanied by a polarity inversion at the mastoid electrodes (Näätänen, 1992, 1995; Schröger, 1998). More recently, it has been shown that the MMN is not only elicited by oddball series consisting of two stimuli, but also can be elicited if the majority of the stimuli carry a “standard” feature (such as an ascending tone sequence), whereas a minority of the stimuli, the “deviants,” do not share this feature. In such a case, the deviants also elicit an MMN (for a review, see Näätänen, Tervaniemi, Sussman, Paavilainen, & Winkler, 2001). Therefore, it is assumed that the MMN is the result of an early scanning mechanism, which compares the incoming information to sensory-memory traces of the preceding auditory stimuli (Näätänen & Winkler, 1999). More important for current purposes is the finding that the MMN is not only sensitive to auditory differences between standard or standards and deviant or deviants, but

also to the linguistic relevance of the difference between standard and deviant. In an already seminal study, Näätänen et al. (1997) used the vowels /e/ and /õ/ as standard and deviant and presented them to Finnish and Estonian listeners. Both vowels are used in Estonian, whereas Finnish only uses an /e/. Consistent with the assumption that the linguistic relevance of the difference between standard and deviant codetermines the size of the MMN, they found a larger MMN for this contrast in Estonian than in Finnish listeners. However, when vowels that are used in both languages were used as standard and deviant (such as /e/, /o/, and /ö/), similar MMNs were observed in both groups (for similar results see Näätänen, 2001; Phillips, 2001).

Mitterer and Blomert (2003) applied the MMN to the study of the perception of assimilated word forms. They presented the same acoustic change—[tœyn] *garden* as standard and [tœym] as deviant—in two contexts. In one context, a participant would hear a sequence such as “... , [tœynbaŋk], [tœynbaŋk], [tœymbaŋk], [tœynbaŋk], ... ,” in which the deviant could arise by assimilation of the [n] in [tœyn] to [m]. In the other context, the context word [baŋk] *bench* was replaced by [stuːl], *chair*. In this case, the deviant is *[tœymstuːl], which could not arise by assimilation of the [n] in [tœyn] to [m]. The results obtained with these stimuli in a passive-oddball task were consistent with the notion of a context-sensitive interpretation of the change of /n/ to [m]: An MMN was observed for the deviant that could not arise from the application of phonological assimilation in the standard (*[tœymstuːl] vs. [tœynstuːl]), whereas no reliable MMN was found for the standard–deviant pair [tœynbaŋk] vs. [tœymbaŋk]. One attractive feature of this design is the inherent control for the acoustic difference between standard and deviant. That is, there is the same amount of acoustic deviancy between the two tokens [tœyn] and [tœym] in both contexts. Hence, the fact that a larger MMN was observed for the pair *[tœymstuːl] versus [tœynstuːl] than for the pair [tœynbaŋk] versus [tœymbaŋk] cannot be attributed to the physical differences between standard and deviant, because the physical differences between both pairs are identical. These results thus reveal a sensitive process of compensation for assimilation: The same segment is interpreted differently depending on the viability of the phonological assimilation.

Although the MMN design does not investigate how lexical access is achieved for assimilated words, it indicates that the deviation from the canonical form is compensated for relatively early, possibly nonlexically, because the MMN in the study of Mitterer and Blomert (2003) arose before the complete words were presented. Although this does not preclude the possibility that lexical processes may still influence the compensation for assimilation (but see Mitterer and Blomert, Experiment 3), it indicates that the compensation is a purely perceptual procedure independent of decision processes. Assuming a relatively early, possibly prelexical, locus of compensation for assimilation does not, however, imply that lexical processes only start after prelexical processes have been completed (see McQueen, Dahan, & Cutler, 2003, for a summary of the evidence against this position). Accordingly, assuming a relatively early process of compensation for assimilation does not exclude the possibility that higher level processes may influence the interpretation of word forms such as [rʌm] as either *rum* or *run*, which has undergone assimilation. In that case, however, such higher level effects seem to be statistically additive to already-established phonological context effects (Gaskell & Marslen-Wilson, 2001).

Therefore, the passive-oddball design introduced by Mitterer and Blomert (2003) may prove useful here. Because listeners neither attend to nor react to the presented speech stimuli

during this task, the “task”—watching a silent movie—in this cross-linguistic design is similar for both native and non-native listeners. This allows us to test and compare the effect of language experience on the perception of assimilated utterances in native and non-native listeners.

2. Experiment 1

In this experiment, we tested the perception of Hungarian liquid assimilation by Dutch listeners. To make this experiment comparable to the study by Mitterer and Blomert (2003), we made use of Dutch words. To these Dutch words, Hungarian liquid assimilation was applied. In one of the words we used, [knaɫroːt] *vivid red*, the assimilation to [knaɫroːt] is viable according to Hungarian phonology. Note, however, that neither this rule nor a similar rule exists in Dutch (Booij, 1995). Therefore, (*)[knaɫroːt] should actually be indicated with a star, as this form is not allowed in Dutch. In the other word we used, [knaɫblau] *vivid blue* the assimilation to *[knaɫblau] is not only forbidden in Dutch, but also in Hungarian.

These words were presented in two oddball series. In both series, the standard contained the unaltered word [knaɫ] and the deviant was [knaɫ]. In one of the series, these targets were followed by [roːt] *red*, in the other by [blau] *blue*. If Dutch listeners treat Hungarian liquid assimilation in a way similar to Dutch nasal place assimilation, the MMN to the viable deviant [knaɫroːt] should be smaller than the MMN to the unviable deviant [knaɫblau]. Note that such a difference cannot be explained by physical stimulus attributes, because the difference between standard and deviant is identical in both cases.

2.1. Method

2.1.1. Participants

Fifteen participants joined this study for a course credit. Participants ranged in age from 18 to 25 years (mean: 20.3 years); 13 participants were right-handed and two were left-handed. All participants were native speakers of Dutch and free of hearing problems. All of the participants spoke English reasonably well, and all but 1 participant spoke at least a second foreign language (either German or French). One participant spoke some Polish; however, none of the Dutch listeners knew any Hungarian. All participants gave their informed consent.

2.1.2. Materials

We recorded a speaker of Dutch, who uses an apical trill for the phoneme /r/. (The majority implementation for a trill in Dutch is the uvular trill; however, apical trills are also observed.) The speaker uttered the words [knaɫroːt] (Engl. “vivid red”) and [knaɫblau] (Engl. “vivid blue”) several times. One of the tokens of [knaɫ] spliced from [knaɫblau] was edited. All splice points were at major positive zero-crossing of the waveforms, preventing clicks and phase shifts in the glottal cycles in the cross-spliced stimuli. Given the similar place of articulation for the apical trill and the apical lateral, the primary cue for the lateral–trill distinction is the presence of amplitude modulation (AM) in the trill (cf. Ladefgèd & Maddieson, 1996). In addition, the resonant characteristics of the lateral and the open period of the trill are slightly

different. To have a model for the editing, a native speaker of Hungarian was recorded pronouncing multimorphemic Hungarian words (see Materials in Experiment 2) that allow liquid assimilation to occur in canonical form and are purposefully assimilated. In the purposefully assimilated utterances, the geminate [r:] consisted out of two or three trill cycles of AM. Given that both accounts of liquid assimilation (see Olsson, 1992, p. 57; Siptár & Törkenczy, 2000, p. 182) described it as a process restricted to casual speech, we decided to use one cycle of AM for the changed [knər], so that the assimilated segment would not be hyperarticulated.

To manipulate both cues that differentiate [r] from [l]—resonant characteristics and AM—the original stimulus was divided in a source and a filter by a linear predictor coefficients analysis (16th order). For the edited sound stimuli, one cycle of AM (20 Hz, 12 dB) was added to the estimated source. This was done by editing the intensity function of the source as estimated by PRAAT (Version 4.0; Boersma & Weenink, 2001). Formant estimations showed a slightly (200 Hz) lower third formant for the resonant period of the trill. Therefore, the third formant was also lowered by this amount. The edited word started to differ from the original at 200 msec. These changes were sufficient for listeners to perceive the edited stimulus with an [r] in a context that cannot trigger assimilation (Mitterer, Csépe, & Blomert, in press). The edited words [knəl] and [knər] were then concatenated with the context words [blau] and [ro:t]. These were spliced from two other utterances. This procedure yielded an unchanged stimulus with a viable context [knəlro:t], an unchanged stimulus with an unviable context [knəlblau], a changed stimulus with an—according to the Hungarian rules—viable context [knərro:t], and a changed stimulus with an unviable context *[knəlblau] (see Fig. 1A). These stimuli were presented over a loudspeaker under control of the ERTS package (BeriSoft Cooperation, Frankfurt, Germany).

To confirm that the edited stimuli are similar to spontaneously produced liquid assimilations, we ran a small production study. We selected the 15 most common /l/-final words, which were to be produced in three different sentence contexts, with a sentence length of about 10 to 20 words. The sentences were taken from the transcribed informal conversations of the Hungarian National Corpus (<http://corpus.nytud.hu/mnsz/>) and slightly adapted to our needs as follows: In two sentence contexts, the /l/-final word was followed by either the case suffix *ra* (“where on”) or the case suffix *ról* (“from where”). In both cases, liquid assimilation can occur. In a third sentence context, the /l/-final word was followed by a nonassimilatory suffix (e.g., *nál*). The three sentence contexts were spread over three lists, and three speakers read one of the lists three times. This gave rise to 90 (3 speakers × 3 repetitions × 10 words) utterances, in which an /l/-final morpheme was followed by an /r/-initial suffix. Analysis of the spectrograms of the utterances revealed that although several utterances showed some weak evidence of assimilation, that is, presence of AM and absence of a liquid-like spectrum, one utterance clearly classified as a liquid assimilation (/hɔlra/ → [hɔrra]). Fig. 2 shows this token, as well as an unassimilated version of the same word by the same speaker. The assimilated token has two cycles of amplitude modulation, and also a slightly lower F3 in the [r] than the [l] in the unassimilated token. Hence, this natural assimilation is not different from the edited speech tokens used in the MMN study, and our stimuli certainly fit very well into the (small) range of assimilations produced.

Unfortunately, due to the low incidence of clear liquid assimilation in this small study, it is not possible to investigate more intricate issues, such as gradual versus categorical assimilation. Moreover, one might also choose to characterize the process as deletion or reduction of

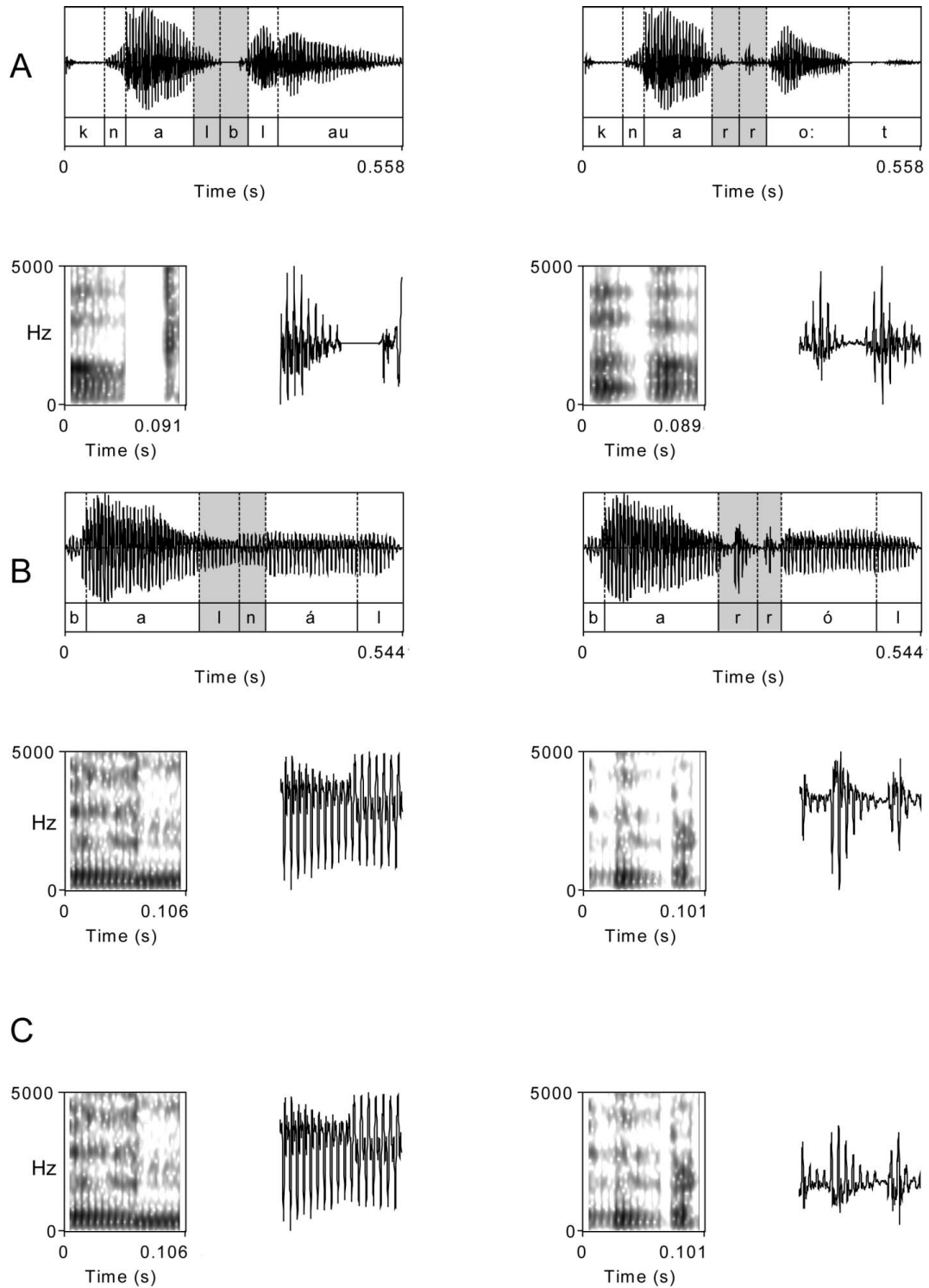


Fig. 1. Stimuli used in Experiment 1 (Panel A), Experiments 2 and 3 (Panel B), and Experiments 4 and 5 (Panel C). In each panel, the oscillogram for the unchanged, viable-context stimulus is shown on the left, and the changed, viable-context stimulus on the right. The oscillogram and spectrograms of the critical highlighted areas are shown below each whole-stimulus oscillogram. In Panel C, the reiteration of the whole stimulus has been left out; see panel B for the whole stimuli.

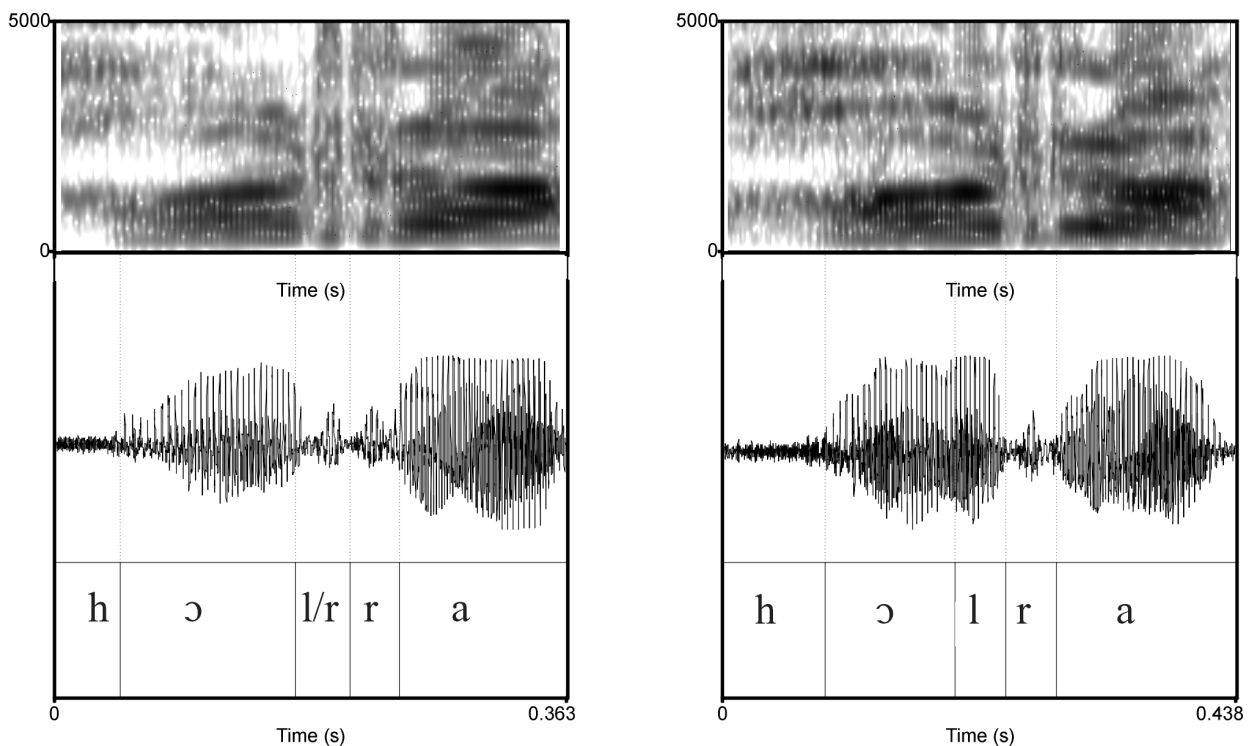


Fig. 2. An example of spontaneous, unintentional liquid assimilation by a native speaker of Hungarian, saying /hɔlra/ → /hɔrra/ to the fish on the left and an unassimilated version by the same speaker on the right.

the final /l/, instead of assimilation. Even in this case, the phonological process occurring in Hungarian does not occur in Dutch, even when casual speech is investigated (cf. Ernestus, 2000). Accordingly, Dutch listeners have no experience with the phonological process under investigation. It should be noted, however, that it is quite likely that the phonological modification of final /l/ occurs more often in natural conversations. Mitterer and Ernestus (2006) investigated /t/ reduction in Dutch and found that in a sentence-production task, less than 1% of the /t/s were severely reduced, whereas in the corpus of spoken Dutch (Corpus Gesproken Nederlands), 8% of all /t/s were severely reduced in read stories, and more than 43% of the /t/s were severely reduced in spontaneous speech. Unfortunately, a comparable corpus of Hungarian speech samples does not exist.

2.1.3. Procedure

After mounting of the electrodes, participants were seated in a comfortable chair and watched a silent video. At the same time, participants heard an oddball series over headphones with 1 stimulus every 1.2 sec. There were four blocks with 805 stimuli each. Each block started with 5 standard stimuli that were not used for the ERP averaging. In addition, every deviant was followed by at least 1 standard stimulus.

Within one recording session, two different oddball series were presented in two blocks each. In one series, the standard ($p = .85$) was [kɲalro:t], and the deviant was the Hungarian-viable alternative pronunciation [kɲarro:t]. In the second series, the standard was [kɲalblau] ($p = .85$) and the deviant was the unviable alternative pronunciation *[kɲarblau]. In half of the recording sessions, the blocks with the viable alternative as deviant were the first

and third blocks, and the blocks with the unviable alternative were the second and fourth blocks. In the other half of the recording sessions, this order was reversed.

2.1.4. Electrophysiological recording and data reduction

The nose-referenced EEG (0.1–27 Hz, sampling rate 256 Hz) was recorded with a 32-channel electrode cap covering frontal, central, temporal, and parietal scalp areas. Blinks and vertical eye movements were monitored with electrodes placed at the sub- and supraorbital ridge of the right eye. Lateral eye movements were monitored by a bipolar montage using two electrodes placed on the right and left external canthus. All electrode impedances (EEG and electrooculogram) were kept below 5 kOhm. Using a 32-channel SynAmp amplifier (Neuroscan, El Paso, Texas) and the SCAN program of the Neuroscan software package (Neurosoft, El Paso, Texas), the brain electric activity was acquired in continuous mode. To generate ERPs, the acquired EEG was sliced into epochs from 52 msec before stimulus onset to 1,000 msec after stimulus onset, band-pass filtered from 1 to 20 Hz, and baseline-corrected from –50 to the point at which standard and deviant started to differ, which was at 200 msec. Artifacts from vertical eye movements were reduced using linear regression. After decorrelation, samples were rejected if the voltage on any channel excluding the vertical eye channel exceeded a value of $175\ \mu\text{V}$. Epochs were then averaged for each stimulus type. The individual ERP averages were obtained from 240 epochs in each condition by using all deviant stimuli and the same number of standard stimuli randomly drawn from all standard stimuli, excluding standards directly following deviants.

For data analysis, a mean-amplitude measure from Fz was used. Mean amplitudes were determined from a 100-msec window with the peak of the grand-average MMN as anchor (cf. Schröger, 1998, p. 135; see also Näätänen, Schröger, Karakas, Tervaniemi, & Paavilainen, 1993; Näätänen et al., 1997; Winkler et al., 1999, for similar methods). This grand-average peak was determined by visual inspection of Fz and mastoid electrodes.

2.1.5. Design

The design entails two independent variables. The first variable is coined Change, with the levels “unchanged” and “changed.” The second variable is coined Context, with the levels “viable” and “unviable.” The unchanged viable stimulus is [kna:lro:t], the changed viable stimulus is [knar:ro:t], the unchanged and changed viable stimuli are [kna:lblau] and [knarblau], respectively. The dependent variable is the mean amplitude on Fz in a 100-msec window around the grand-average peak of the MMN.

2.2. Results and discussion

Fig. 3 shows the grand-average waveforms from Fz, M1, and M2 for all four conditions. The MMN was more pronounced in the unviable-context condition and peaked at 388 msec, which is about 180 msec after the onset of the change. The mean amplitudes (see Table 1) were subjected to an analysis of variance (ANOVA) with Change and Context as predictors, which revealed a significant effect of Change, $F(1, 14) = 20.27, p < .001$, and a significant effect of Context, $F(1, 14) = 8.71, p < .025$. More important, there was also a significant interaction between Context and Change, $F(1, 14) = 7.86, p < .025$. Subsequent *t* tests showed that the effect of Change (hence, the MMN) was significant in the unviable context, $-0.97\ \mu\text{V}, t(14) = 4.40, p$

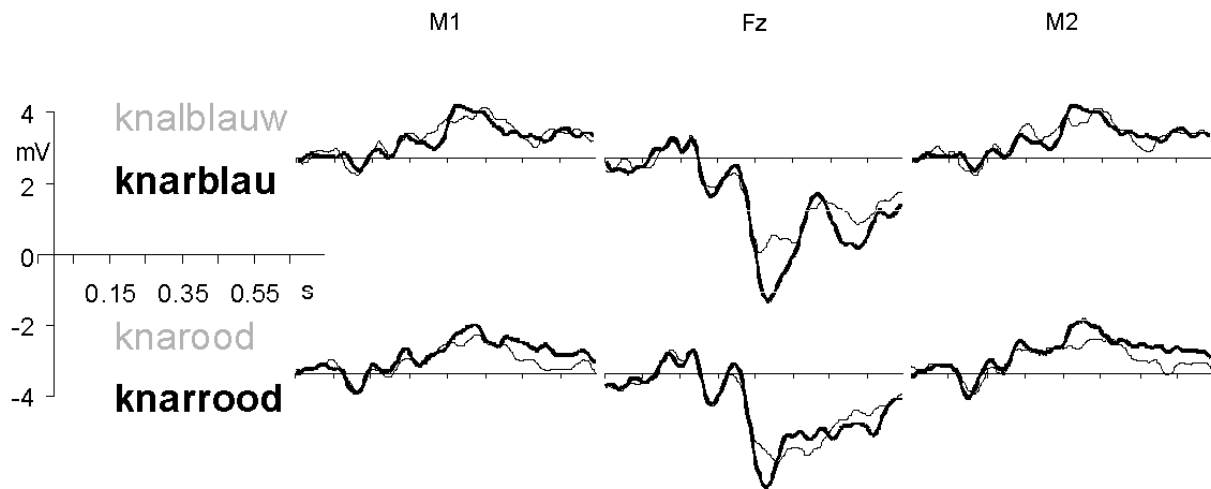


Fig. 3. ERPs from the electrodes Fz, M1, and M2 in Experiment 1. The thick lines show the ERPs for the deviant stimuli, the thin lines the ERPs for the standard stimuli.

< .005, but constituted only a marginally significant effect in the viable context, $-0.31 \mu\text{V}$, $t(14) = 2.02$, $p < 0.1$.

In this experiment, the MMN was smaller for the deviant, which was a viable assimilation of the standard—according to Hungarian, but not Dutch phonology. The MMN elicited by $(*)[\text{knarro:t}]$ with $[\text{kndlro:t}]$ as standard was smaller than the MMN elicited by $*(\text{knarblau})$ with $[\text{kndlblau}]$ as standard. The pattern of this context sensitivity was as expected. That is, the MMN was smaller if the deviant stimulus was a viable alternative pronunciation, according to the Hungarian assimilation rule, than in a case where the same acoustic change was not a viable change. This seems to indicate that compensation for assimilation is indeed independent of language experience: Dutch listeners treat Hungarian liquid assimilation similarly to Dutch nasal place assimilations.

One possibility how the results may nevertheless be explained in terms of specific language experience is that the Dutch listeners may have experience with [l]-deletion in casual speech. This would change the pronunciation of $/\text{kndlro:t}/$ to $[\text{knarro:t}]$, which is similar to the form $[\text{knarro:t}]$, which arises by liquid assimilation. Ernestus (2000, pp. 112–125) presented an overview of segment deletions that occur in Dutch. She found that the segments $/t/$, $/r/$, $/n/$, and $/k/$ can be deleted in Dutch generatively, whereas the deletion of $/l/$ seems to be limited to a few high-frequency words such as $/\text{als}/$, *like*, or $/\text{naturlək}/$ *naturally*, which are two to three orders of magnitude more frequent than the word *knal*. It seems therefore unlikely that Dutch listeners are familiar with $/l/$ -deletion as a general process, which would lead them to treat the forms $[\text{kndlro:t}]$ and $[\text{knarro:t}]$ as similar, because of experience with the form $[\text{knarro:t}]$.

It is difficult to explain this results in terms of top–down effects in speech perception (e.g., Ganong, 1980; Magnuson, McMurray, Tannenhaus, & Aslin, 2003). Top–down connections from the lexicon to a prelexical level could have lead to the correction of the prelexical representation of $[\text{knarro:t}]$ to conform with the existing word $[\text{kndlro:t}]$. This explanation would then also apply to the stimulus $[\text{knarblau}]$. It is hence difficult to explain the statistical interaction of Context and Change in our experiment in terms of an interaction between prelexical and lexical processing.

These results add to the existing literature establishing context sensitivity in the perception of phonologically assimilated utterances (Coenen, Zwitserlood, & Bölte, 2001; Gaskell & Marslen-Wilson, 1996, 1998, 2001; Gow, 2002, 2003; Mitterer & Blomert, 2003). Given this prelexical context sensitivity, it seems unnecessary to invoke the notion of phonologically underspecified lexical representations of assimilated features. Lahiri and coworkers (Lahiri & Marslen-Wilson, 1991; Lahiri & Reetz, 2002) argued that assimilated words are recognized independently of phonological context because lexical representations are underspecified for features that assimilate. Hence, in this framework, the Dutch word [tœyn] *garden* is represented lexically as /tœy + nasal/, which fits the canonical form [tœyn] as well as the assimilated form [tœym]. However, our results show that assimilated utterances are recognized depending on the context.

This experiment therefore shows that Dutch listeners compensate for Hungarian liquid assimilation just as they do for Dutch nasal place assimilation. However, what we have not yet shown is that Hungarian listeners compensate for their native liquid assimilation in a fashion similar to Dutch listeners. This was the purpose of Experiment 2.

3. Experiment 2

In this experiment, we evaluated whether the MMN to phonological assimilations is context dependent when Hungarian words are presented to Hungarian listeners. Although this means that the stimuli are different than in Experiment 1, they are conceptually similar: In both experiments the stimuli are words of the native language of the listener presented with a familiar accent.

The target word was Hungarian word *bal* [bɔl] *left*. This word was presented in either the canonical form or changed as [bɔr]. Both changed and unchanged forms could be followed by either the case suffix [ro:l] (delative, similar to *from*) or the case suffix [na:l] (adessive, similar to *at*). According to the rule of Hungarian liquid assimilation (see Olsson, 1992, p. 57; Siptár & Törkenczy, 2000, p. 182), the change from /l/ to /r/ is a possible assimilation before the delative suffix [ro:l] but not before the adessive suffix [na:l]. Thus, based on the assumption that there is a context-sensitive compensation mechanism, we expect the MMN to the pair [bɔlro:l]–[bɔrrɔ:l] to be smaller than the MMN to the pair [bɔlna:l]–*[bɔrna:l]. It should also be noted that the Hungarian canonical form [bɔl] may be perceived as the Dutch word /bɔl/*globe*—even though the Hungarian vowel is not a good exemplar of the Dutch one. However, [bɔr] is not a Dutch word, and none of the four two-syllable utterances can be parsed as a Dutch word or a combination of Dutch words.

3.1. Method

3.1.1. Participants

Twelve students (6 men, 6 women) from different Budapest universities participated in the study. They were paid for participation. Participants ranged in age from 16 to 28 years (mean: 20.5 years); 10 participants were right-handed and 2 were left-handed. All participants gave their informed consent, completed a handedness questionnaire, and were screened for hearing problems. They were all native speakers of Hungarian.

3.1.2. Materials

A female native speaker of Hungarian was recorded uttering multiple tokens of the canonical form [bɔlna:l] (Engl. “at the left”), with an unviable variant *[bɔrna:l], and the canonical form [bɔlro:l] (Engl. “from the left”), with a viable alternative [bɔrro:l]. The Hungarian context suffixes of the cases called delative (from where?) and adessive (where at?) were chosen, because they are phonetically quite similar. Both start and end with a sonorant and contain a long vowel. Any differences that these stimuli may cause in the perception of the preceding segments can thus not be attributed to gross acoustic differences (overall amplitude, presence of voicing) between the context sounds.

Sample frequency for recording the natural utterances was 22050 Hz. Recordings were band-pass filtered from 130 to 8000 Hz. (minimal f0 was 150 Hz.) The syllable [bɔl] was spliced from an utterance of the compound [bɔlna:l] and edited in the same way as the syllable [knɔl] in Experiment 1 to yield the form [bɔr]. Filtering the edited source with the edited filter yielded a stimulus that was perceived as an apical trill in the behavioral study of Mitterer et al. (in press). To convey a casual speaking style, prevoicing was shortened from 40 msec to 27 msec by cutting two complete cycles of prevoicing in all stimuli starting with a voiced labial stop. These stimuli had a length of 544 msec.

The context case suffixes [na:l] (Engl. *at* in answer to the question “where?” as in “at your left”) and [ro:l] (Engl. *from*) were spliced from other utterances, equalized in overall energy, and concatenated with the Hungarian word and nonword stimuli. The amplitude relations of word and nonword to the case suffixes were edited so that they emulated the amplitude relations of the first (always stressed) syllable to the second syllable in the natural utterances. Fig. 1B shows the stimulus [bɔlna:l] in which the first syllable has no AM, and the second syllable is the case suffix [bɔrna:l] that does not allow the assimilation [l] to [r]. These stimuli were presented to the participants over headphones using the Presentation software (Neurobehavioral Systems, Albany, California).

3.1.3. Procedure, electrophysiological recording, and data reduction

The EEG was recorded over 20 (Fz, Cz, Pz, Oz, F3, F4, F7, F8, C3, C4, T3, T4, T5, T6, P3, P4, O1, O2, M1, M2) electrodes linked to the nose as reference. Otherwise, procedure, electrophysiological recording, and data reduction were the same as in the previous experiments. Note that this leads to a baseline for the ERP from –50 to 220 msec, because this is the time point at which standard and deviant start to differ.

3.1.4. Design

The design entails two independent variables. The first variable is coined Change, with the levels “unchanged” and “changed.” The second variable is coined Context, with the levels “viable” and “unviable.” The dependent variable is the mean amplitude on Fz in a 100-msec window around the grand-average peak of the MMN.

3.2. Results and discussion

Fig. 4 shows the grand-average waveforms from Fz, M1, and M2 for all four conditions. There is a well-defined MMN that reverses polarity at the left (M1) and right (M2) mastoid electrodes. The polarity reversal at the mastoid electrodes is usually taken as a signature of the

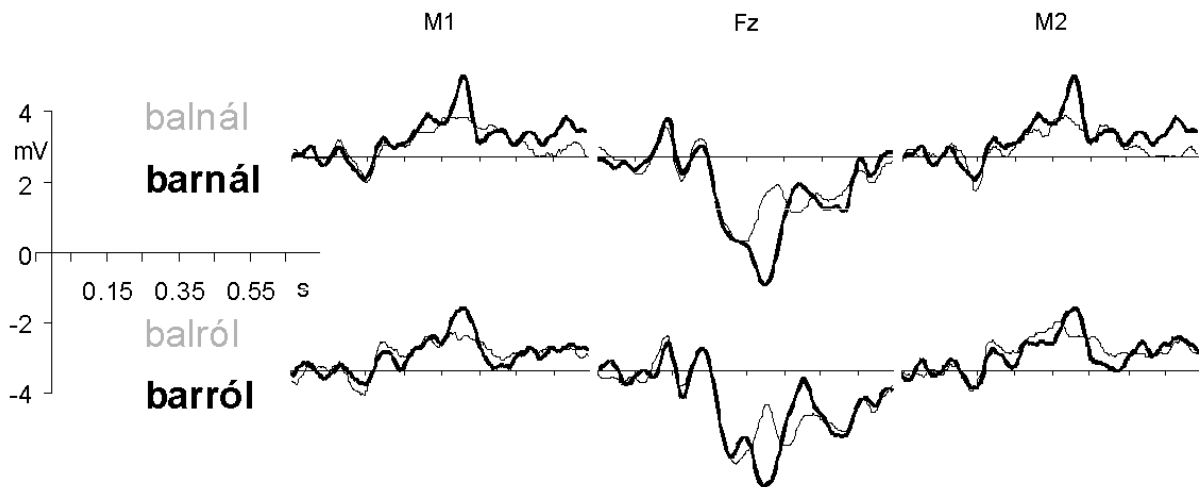


Fig. 4. ERPs from the electrodes Fz, M1, and M2 in Experiment 2. The thick lines show the ERPs for the deviant stimuli, the thin lines the ERPs for the standard stimuli.

MMN (cf. Schröger, 1998). The MMN peak is at 404 msec poststimulus onset, which is again approximately 180 msec after onset of the acoustic mismatch between standard and deviant. The mean amplitudes in a window of 354–454 msec (see Table 1) were subjected to a two-factorial repeated measures ANOVA with the factors Change and Context. The ANOVA revealed a significant effect of Change, $F(1, 11) = 88.97, p < .001$, indicating that the deviants had a more negative ERP. All other effects, including the crucial interaction of Context and Change, $F(1, 11) = 2.97, p > .1$, failed to reach significance (Context: $F < 1$).

The data show an MMN that has a peak around 180 msec after the gradual onset of the acoustic mismatch between standard and deviant. This negative component reverses polarity at the mastoid electrodes, which fosters its interpretation as an MMN (cf. Schröger, 1998). The MMN observed here occurs in a time frame similar to Experiment 1. However, there is no context sensitivity of the MMN. This contradicts our initial hypothesis. We expected the MMN to the [bɔlrɔ:l]–[bɔrrɔ:l] pair to be smaller than the MMN to the [bɔlna:l]–*[bɔrna:l] pair. However, the two MMNs are of similar size.

This is rather surprising, given the finding of context sensitivity in Experiment 1. There are two differences between these two experiments that might explain the difference in results: the stimuli used and the language background of the listeners. In the next experiment we tested whether the difference in language background (Dutch vs. Hungarian listeners) might explain the difference in results.

4. Experiment 3

In this experiment, we probed the preattentive perception of the phonologically changed Hungarian utterances by Dutch listeners. The same utterances as in Experiment 2 were used. None of the utterances are meaningful in Dutch. These stimuli were used in oddball series in which the canonical word is the standard and the (in Hungarian) viable or unviable alternative is the deviant.

4.1. Method

4.1.1. Participants

Twelve students (11 women, 1 man) from the University of Maastricht participated in the nonword part of the experiment. They were paid for participation. Participants ranged in age from 18 to 25 years (mean: 20.3 years); 10 participants were right-handed, two were left-handed.

All participants were native speakers of Dutch and free of hearing problems. All of the participants spoke English reasonably well, and all participants spoke at least a second foreign language (either German or French). One participant spoke some Polish; however, none of the Dutch listeners knew any Hungarian. All participants gave their informed consent.

4.1.2. Materials and procedure

Materials and Procedure were the same as in Experiment 2.

4.1.3. Electrophysiological recording and data reduction

The nose-referenced EEG (0.1–125 Hz, sampling rate 250 Hz) was recorded with a 32-channel electrode cap covering frontal, central, temporal, and parietal scalp areas. An electrode on the nose served as a reference. EOG was measured as in Experiment 1. ERPs were generated in the same way as in Experiment 1.

4.1.4. Design

The design was the same as in the previous experiments, with two independent variables with two levels (Change: standard vs. deviant, and Context: viable and unviable). The independent variable was the mean amplitude on Fz in a 100-msec window around the grand-average peak of the MMN.

4.2. Results and discussion

Fig. 5 shows the grand-average waveforms from Fz, M1, and M2 for all four conditions. The ERPs reveal a well-defined MMN that reverses polarity at the mastoid electrodes. The MMN peak is at 400 msec poststimulus onset. This is approximately 175 msec after onset of the deviance, and is quite similar to the 180 msec observed in the Hungarian sample. The mean amplitudes in the window 350–450 msec (see Table 1) were subjected to a repeated measures ANOVA with Change and Context as predictors.

The ANOVA revealed a significant effect of Change, $F(1, 11) = 18.80, p < .005$. However, neither the effect of Context ($F < 1$) nor the interaction between Context and Change, $F(1, 11) = 1.96, p > .15$, were significant. These results are rather similar to the results obtained in Experiment 2. The MMN has a very similar latency, and again we fail to find an effect of Context on the MMN.

This result rules out the possibility that the difference in language background is responsible for the difference in results between Experiments 1 and 2. Instead, the different stimulus material may give rise to the difference in results. Fig. 1 shows the stimuli used in this study. A detailed comparison of the deviant stimuli reveals that the peak of the AM in the deviants—the

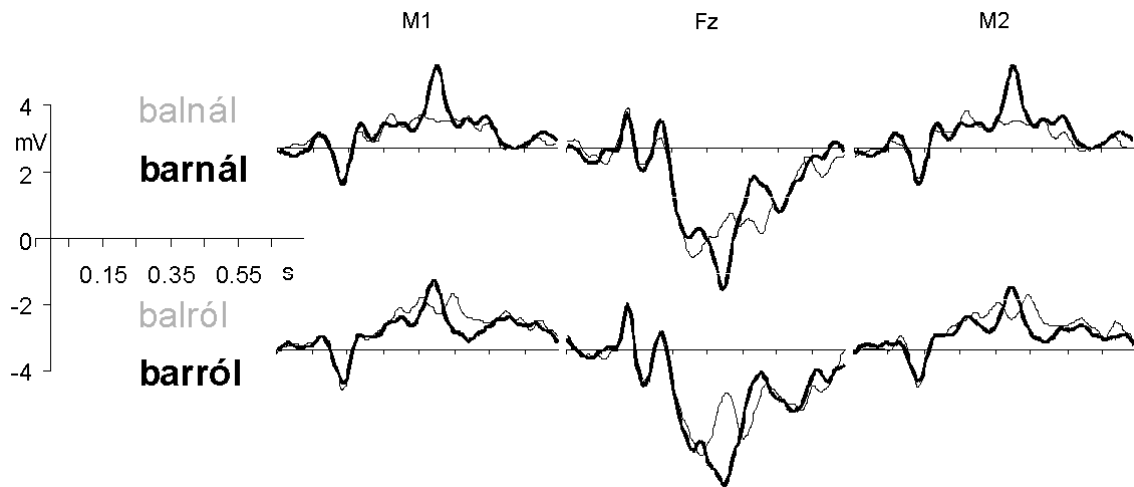


Fig. 5. ERPs from the electrodes Fz, M1, and M2 in Experiment 3. The thick lines show the ERPs for the deviant stimuli, the thin lines the ERPs for the standard stimuli.

first peak—was larger than the AM peak in the viable-context stimulus—the second AM peak—in the case of the Hungarian stimuli (Fig. 1A), but not in the Dutch stimuli (Fig. 1B). (This difference arose despite the identical manipulation procedure from the lower amplitude of the [l] in the model [knəl] than in the [l] in the model [bəl].) Therefore, the difference in results between Experiment 1 on the one hand and Experiments 2 and 3 on the other might be due to the relation of the AM peaks in deviant and context sounds. This assertion is easily tested: A similar effect should be observed when the Hungarian word is altered to have similar acoustic properties as the Dutch word.

Such a prediction is not as far-fetched as it may seem. Gow (2002, 2003) essentially showed that an English phrase such as [raip beriz] is interpreted as “ripe berries” if it contains a good example of a /p/, but as “right berries” if it contains a poorer example of /p/. It is also important to note that such an effect of acoustic detail is highly functional. Gow also showed that assimilation may give rise to phrases such as [raip beriz] when “right berries” was intended, but the /p/ is less strong than when the speaker intended to say “ripe berries” (see also Nolan, 1992). Hence, it makes sense for the perception system to treat a weak /p/ as possibly assimilated—and perceive it as a possible /t/ if followed by /b/—but to perceive a strong /p/ faithfully, independent of what the phonological context is. If this is translated to our stimuli, strong examples of an /r/ such as used in Experiments 2 and 3 may not be perceived in a context-sensitive way, whereas a weak /r/, such as the token in Experiment 1, will be perceived in a context-sensitive way.

5. Experiment 4

In this experiment, we tested whether a context-sensitive MMN—larger in the unviable context than in the viable context—can also be obtained when altered versions of the Hungarian word used in Experiments 2 and 3 were presented to Dutch listeners.

To this end, we used the Hungarian word [bɔl] in its original form as standard. As deviant, a less drastically different form with a lower AM peak (depth = 9 dB; cf. Fig. 1A with Fig. 1C) was used. As in the previous experiments, these stimuli were presented in one context that allows the assimilation [...ro:l] and one that does not allow the assimilation [...na:l]. In the study of Mitterer et al. (in press), we tested the perception of a wider range of AMs in these stimuli. The stimulus used in this experiment was still perceived predominantly as [bɔr].

5.1. Method

5.1.1. Participants

Ten students (8 women, 2 men) from the University of Maastricht participated in the experiment. They were paid for participation. Participants ranged in age from 19 to 26 years (mean: 21.0 years); 9 participants were right-handed, 1 was left-handed. All participants were native speakers of Dutch and free of hearing problems. All of the participants spoke English reasonably well, and all but 2 participants spoke at least a second foreign language (either German or French). No other languages were spoken by any participant. All participants gave their informed consent.

5.1.2. Materials

The same standard stimuli as in Experiment 2 were used. New deviant stimuli were created by choosing an AM depth that led to similar maximal amplitudes of the AM in the target and the AM in the context stimulus (see Fig. 1C). These Hungarian stimuli were nonwords for the Dutch listeners.

5.1.3. Procedure, electrophysiological recording, and data reduction

Procedure, electrophysiological recording, and data reduction were the same as in Experiment 1.

5.1.4. Design

The design entails two independent variables. The first variable is coined Change, with the levels “unchanged” and “changed.” The second variable is coined Context, with the levels “viable” and “unviable.” The dependent variable is the mean amplitude on Fz in a 100-msec window around the grand-average peak of the MMN.

5.2. Results and discussion

Fig. 6 shows grand-average ERPs for all four conditions. There is a clear peak of the MMN at Fz with a peak at 404 msec accompanied by a positive peak at the mastoid electrodes. This replicated the latency value of the MMN found in the earlier experiments. The mean amplitudes in the 354 to 454-msec window (see Table 1) were subjected to a repeated measures ANOVA with Change and Context as predictors. The ANOVA revealed a marginally significant effect of Change, $F(1, 9) = 4.06, p < 0.1$, and no significant effect of Context, $F < 1$. However, the interaction between both predictors was significant, $F(1, 9) = 10.96, p < .01$. Consecutive *t* tests showed that the effect of change—and thus the MMN—was significant for the

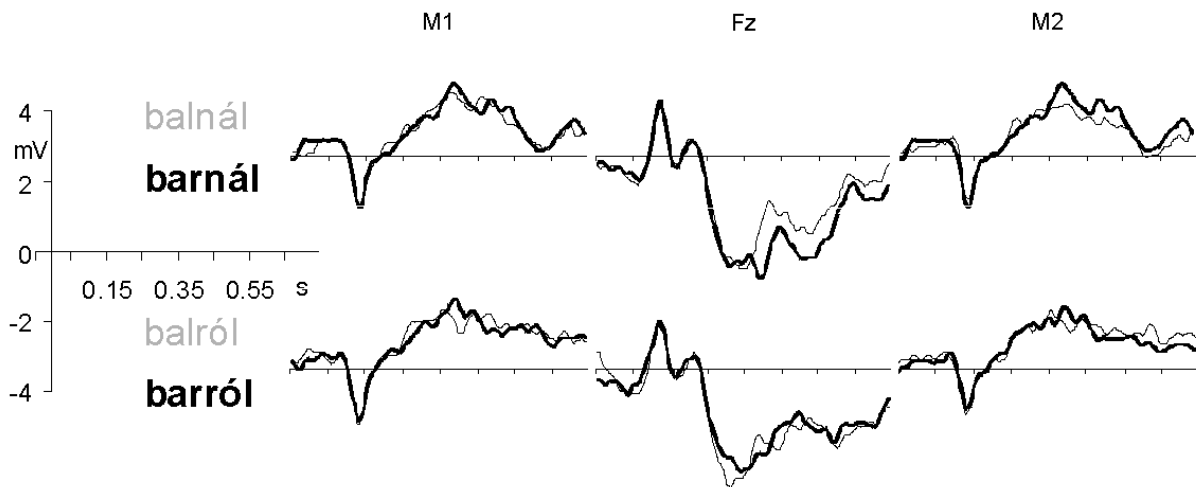


Fig. 6. ERPs from the electrodes Fz, M1, and M2 in Experiment 4. The thick lines show the ERPs for the deviant stimuli, the thin lines the ERPs for the standard stimuli.

unviable-context condition, $t(9) = 2.72$, $p < .025$, but not for the viable-context condition ($t < 1$; see also Table 1).

This outcome clarified the reason for the difference in results between Experiment 1 on the one hand and Experiments 2 and 3 on the other hand: The acoustic details of the stimuli used determine whether a context-sensitive MMN can be observed. The question now arises whether this acoustic detail would also explain the absence of context sensitivity in the Hungarian sample in Experiment 1. To this end, we reran Experiment 4, but now with Hungarian participants.

Experiment 4 also deals with a potential alternative explanation of Experiment 1. For Experiment 1, it may be claimed that it is not the viable context that suppresses the perception of the /l-/r/ distinction, but rather the unviable context that facilitates it. In Experiment 1, the viable context was the syllable [blau], in which the initial stop is dissimilar to the sonorants /l/ and /r/, which may have increased the perceptibility of the /l-/r/ contrast in comparison to the—also sonorant—viable context [rot]. In this experiment, both the viable and the unviable context started with sonorant consonants (/r/ and /n/), but still a context effect is observed. Admittedly, the comparison between Experiment 1 and Experiment 4 is confounded with the listeners' familiarity with the phonology of the tokens as well as their lexical status. However, Mitterer and Blomert (2003; comparison of Experiments 1 and 3) found that both variables seem to matter little in compensation for assimilation.

6. Experiment 5

In Experiments 2 and 3, similar results were obtained for both Hungarian and Dutch listeners if identical stimuli are used. In this experiment, we tested whether this also holds for the Hungarian stimuli that induced a context-sensitive MMN when presented to Dutch listeners. These stimuli were now presented to Hungarian listeners.

6.1. Method

6.1.1. Participants

Ten students (all women) from different Budapest universities participated in the experiment. They were paid for participation. Participants ranged in age from 18 to 29 years (mean: 22.4 years); All participants gave their informed consent, completed a handedness questionnaire, and were screened for hearing problems. They were all right-handed native speakers of Hungarian.

6.1.2. Materials and procedure

The same stimuli and procedure as in Experiment 4 were used, which were words for the Hungarian listeners.

6.1.3. Electrophysiological recording and data reduction

Electrophysiological recording and data reduction were the same as in Experiment 1.

6.1.4. Design

The design was the same as in the previous experiments, with two 2-level independent variables (Context: viable and unviable, Change: standard and deviant). The dependent variable was the mean amplitude on Fz in a 100-msec window around the grand-average peak of the MMN.

6.2. Results and discussion

Fig. 7 shows the grand averages for all four conditions. There is a clear peak of the MMN at Fz with a peak at 396 msec accompanied by a positive peak at the mastoid electrodes. This replicated the latency value of the MMN found in the earlier experiments. The mean amplitudes in the 346 to 446 msec window (see Table 1) were subjected to a repeated measures ANOVA with

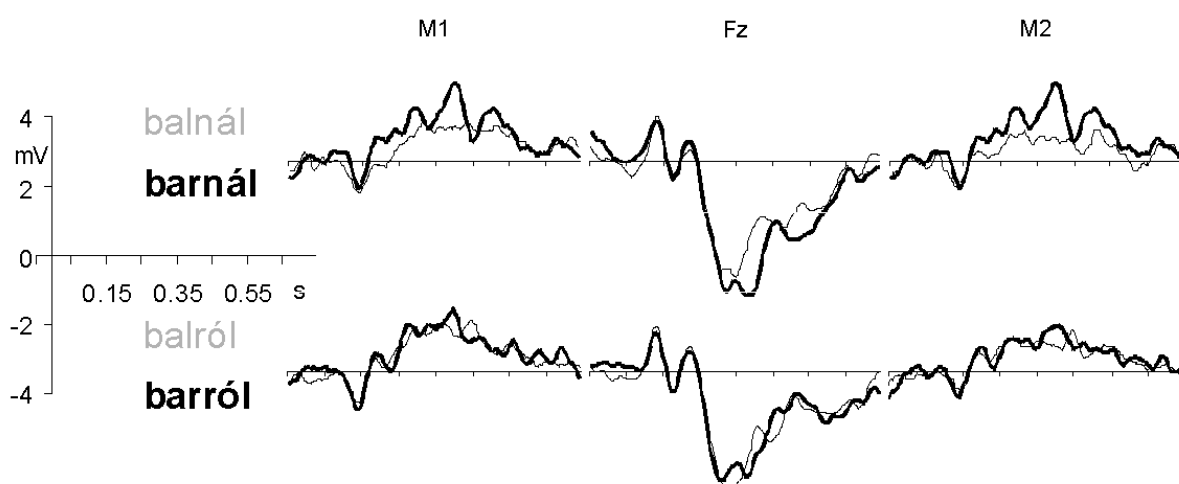


Fig. 7. ERPs from the electrodes Fz, M1, and M2 in Experiment 5. The thick lines show the ERPs for the deviant stimuli, the thin lines the ERPs for the standard stimuli.

Table 1
MMN parameters in Experiments 1 to 5

	Experiment				
	1	2	3	4	5
Stimuli	kna(l/r)	ba(l/r _s)	ba(l/r _s)	ba(l/r _w)	ba(l/r _w)
Participants	Dutch	Hungarian	Dutch	Dutch	Hungarian
Window	338–438	354–454	350–450	354–454	346–446
Voltages in mV per condition					
Change/Context					
–/viable	–1.80	–1.16	–1.78	–1.87	–2.13
+/viable	–2.12	–2.58	–2.99	–1.87	–2.04
–/unviable	–1.95	–1.46	–2.02	–1.67	–1.96
+/unviable	–3.02	–2.33	–2.78	–2.41	–2.86
Effects					
Change	$p < .001$	$p < .001$	$p < .005$	$p < .1$	$p < .05$
Context	$p < .025$	<i>ns</i>	<i>ns</i>	<i>ns</i>	$p < .1$
Interaction	$p < .025$	<i>ns</i>	<i>ns</i>	$p < .01$	$p < .01$

Note. The column stimuli indicate the standards and deviants used in the different experiments. r_s and r_w refer to strong [r] and weak [r] deviants.

Change and Context as predictors. The ANOVA revealed a significant effect of Change, $F(1, 9) = 10.52$, $p < .05$, and a marginally significant effect of Context, $F(1, 9) = 4.30$, $p < .1$. However, the interaction between both predictors was significant, $F(1, 9) = 12.38$, $p < .01$. Consecutive t tests showed that the effect of change—and thus the MMN—was significant for the unviable-context condition, $t(9) = 3.80$, $p < .01$, but not for the viable-context condition ($t < 1$).

This replicated the results of Experiments 1 and 4. It shows that the difference in stimulus material is crucial to get or not get results indicative of compensation for assimilation: only if the changed, and hence possibly assimilated, phoneme is a weak example of the category, does compensation for assimilation occur. Table 1 summarizes the results from all experiments

7. General discussion

In this study, we investigated whether the process of compensation for assimilation depends on experience with a given assimilation rule in one's native language. Utterances in which Hungarian liquid assimilation was applied either viably or unviably were presented to Dutch and Hungarian listeners. Using a passive-listening task with an oddball series, we expected the MMN to be smaller for a standard–deviant pair consisting of a canonical form (e.g. [bɔlro:l], “from the left”) and a viable alternative pronunciation (e.g. [bɔrrɔ:l]) than the MMN for a standard–deviant pair consisting of a canonical form (e.g. [bɔlna:l], “at the left”) and an unviable alternative pronunciation (e.g. *[bɔrna:l]). However, in the initial three experiments such a result was only obtained in Experiment 1, in which Dutch listeners were confronted with Dutch stimuli. Detailed examination of the stimulus material showed that this stan-

dard–deviant pair differed acoustically from the Hungarian stimuli, which failed to reveal a context effect in Experiments 2 and 3. In particular, the AM that introduced the acoustic mismatch between standard and deviant was larger than the AM in the context sound for the Hungarian but not for the Dutch stimuli. Two additional experiments showed that this acoustic detail was critical. Both Dutch and Hungarian listeners showed a context-sensitive MMN with Hungarian stimuli when the AM peak in the deviant was of a size similar to the AM peak of the context (see Table 1).

These results first of all replicated the results obtained by Mitterer and Blomert (2003) with nasal place assimilation. They show that an acoustically identical difference between standard and deviant may give rise to a different MMN depending on the context: If the deviant is a viable assimilated pronunciation of the standard, the MMN is smaller than in cases where the deviant is not a viable assimilated pronunciation of the standard. These results, however, add a complication to this general statement. The finding of a context-sensitive MMN being larger for the /l/-to-/r/ change in the unviable than in the viable context was contingent on the acoustic details of the deviant stimulus: A context-sensitive MMN was only found when the /r/ in the deviant stimulus was a comparably weak /r/. This result replicates an earlier finding on compensation for assimilation. Gow (2002, 2003) found that strong nonalveolar stops, as in [rɪp beriz] produced from intended “ripe berries,” are not subject to context effects, whereas weak nonalveolar stops, as in [rɪp beriz] produced from intended “right berries,” are. It is important to note that this result is not a mere replication of Gow (2002, 2003). Although he obtained results showing the importance of phonetic detail using a cross-modal priming task, we were able to show an effect of phonetic detail using automatic brain responses. Such converging evidence with vastly different methodologies strengthens the claim that phonetic detail is important in the perception of assimilated utterances. Moreover, by replicating Mitterer and Blomert’s result, this study underscores the potential of the MMN to investigate context effects in speech perception, which so far have mainly been investigated with identification and discrimination tasks.

The main objective of this study was to examine whether compensation for assimilation can occur without experience with a given assimilation rule. Experiments 1 and 4 showed that Dutch listeners can compensate for Hungarian liquid assimilation, despite the fact that no similar lenition of /l/ occurs in Dutch. Moreover, the comparison of Experiments 2 and 3 on the one hand and Experiments 4 and 5 on the other shows that Dutch listeners do not differ from Hungarian listeners in how possible assimilated utterances are perceived. Both listener groups compensate for Hungarian liquid assimilation, but compensation depends on the phonetic details of the assimilated segment.

How can these results be united with the earlier evidence that experience with an assimilation rule does influence the perception of assimilated utterances? For the studies of Otake et al. (1996) and Weber (2001), we noted that the effect of experience with an assimilation rule was confounded with a difference in phonemic inventory of the languages investigated. However, Darcy (2002) showed that compensation for a French voicing assimilation by English listeners depends on the level of expertise of the English listeners with the French language. We see two ways to explain the results of Darcy and our own results coherently. One possibility is that language experience enhances language-independent auditory biases (cf. Coady, Kluender, & Rhode, 2003, for a similar view). Another possibility would be to argue that our results only in-

dicating that the difference between /l/ and /r/ is difficult to perceive in front of another /r/. Indeed, our results do not show that Dutch and Hungarian listeners perceive the original and changed stimuli as ending on /l/ in the viable context. An alternative possibility would be that they perceive both stimuli in the viable context as ambiguous, that is, they are unsure whether they heard /l/ or /r/. Indeed, behavioral evidence suggests the latter possibility (Mitterer et al., in press). In this case, specific language experience provides listeners with additional means to infer the identity of the ambiguous segment, such as lexical knowledge and statistical regularities. Given the fact that Darcy used a word-monitoring task, participants' performance probably reflected the merged information from prelexical and lexical mechanisms. This is not to say that the influence of higher level processes in the compensation of assimilation is artifactual. Instead, this explanation indicates how a language-independent compensation effect could be amplified by language-specific knowledge.

These data show that compensation for assimilation can be achieved—at the very least partly—without experience with an assimilation rule. This is consistent with the notion that assimilation rules are shaped perceptually, and hence the listener does not need to adapt to assimilation rules, because assimilation rules are adapted to the needs of the listener (Hume & Johnson, 2001; Hura et al., 1992; Kohler, 1990). Our results bear out the similarity principle as a basic constraint on phonological assimilations as put forth by Steriade (2001, p. 222): “The likelihood that a lexical representation R will be realized as modified R' is a function of the perceived similarity between R and R'.” There are three possibilities how this parity between a speaker's modifications and a listener's perceived similarity could arise. First of all, it could be the consequence of an innate language module that governs both perception and production, as proposed by motor theory (Lieberman, 1996). Second, speakers may possess an implicit knowledge of the listener's perceptual abilities and lexical-access processes (Boersma, 1998; Lindblom, 1990; Steriade, 2001). Both possibilities essentially put knowledge about the listener into the speaker's head, which is a rather strong assumption. A third possibility is more parsimonious: The parity between speakers and listeners may be a consequence of an evolutionary process of the language system itself, in which ease of articulation and ease of perception constitute evolutionary pressures (e.g., de Boer, 2000). Assimilations then arise as mutations as speakers try to reduce articulatory effort, but perceptually salient assimilations hinder communication and hence do not survive. Accordingly, the speaker's knowledge of the listener's abilities does not have to reside in the speaker's head, but rather in the culturally transmitted language system. This proposal ascribes the parity between speakers and listeners to memetic, cultural evolution, rather than to genetic evolution, as motor theory does. Dawkins (1989, p. 189) noted that “language seems to ‘evolve’ by nongenetic means, and at a rate which is orders of magnitude faster than genetic evolution.” Such a cultural evolution model provides a basis for our finding that the perception of assimilated forms does not depend on language-specific learning processes.

One theoretical argument may be leveled against the assumption that compensation for assimilation is largely learning independent. If all language users are compensating for all assimilation rules via low-level auditory mechanisms, significant lexical ambiguity seems to arise. Native speakers of French, for instance, would have to treat labial stops (/p/ or /b/) as potentially alveolar (/t/ or /d/) leading to ambiguity in the phrase *la nappe blanche* “the white tablecloth,” because *nappe* is also a word in French. If compensation for assimilation would apply here—based on auditory processing—unnecessary ambiguity would ensue, because the French language does not

apply place assimilation for final stops. Such an argument, however, overlooks the fact that the phonetic implementation of phonemes varies considerably over languages, and if compensation is based on low-level, possibly auditory processes, these phonetic implementation differences will matter. So if low-level auditory processes compensate for place assimilation applying to the unreleased stops of English, there is no need to assume that the same auditory processes would lead to compensation in the acoustically rather different released stops that occur in French. Accordingly, the assumption that compensation based on auditory processing would unduly increase ambiguity in spoken-word recognition is not required.

Our results show that specific language experience is not required for compensation for assimilation at a prelexical level. In this respect compensation for assimilation seems similar to compensation for coarticulation in liquid–stop sequences (Fowler, Best, & McRoberts, 1990; Mann, 1986). These language-independent effects may be due to basic auditory processes (Lotto, Kluender, & Holt, 1997). But an alternative interpretation is also possible in terms of gestural theories of speech perception such as motor theory (Libermann, 1996) or the theory of direct perception of speech gestures (Fowler, 1996). On theoretic grounds, we argued against motor theory. Nevertheless, both gestural theories are able to account for the context effects in the perception of assimilation utterance. Gestural theories assume that an assimilated segment, being a weak exemplar of a category, is perceived as carrying two closing gestures: In the case of the phrase [raip beriz], the listener perceives an alveolar and a labial closing gesture. The context also carries a labial closing gesture (for the /b/ in [beriz]), and this allows the listener to parse the labial closing gesture from the [raip] so that it is perceived as [rait] (see, e.g., the account of the perception of coarticulatory nasalization in Fowler & Brown, 2000). A similar account, although not in articulatory terms, was proposed by Gow (2001, 2002, 2003). He also claimed that the context effect arises at a language-specific phonological feature-cue level. That is, feature cues are assumed to be grouped by Gestalt principles as laid out in Bregman (1990). Both an auditory interpretation as well as a more speech-specific but learning-independent account can accommodate these data. These findings, therefore, do not allow us to distinguish auditory and other learning-independent accounts such as the feature-parsing account. Classically, this can be done by either searching for similar effects with nonspeech materials or by presenting speech material to nonhuman species. This remains to be done before a more definite conclusion about the locus of the compensation-for-assimilation mechanism can be drawn. However, these results indicate that general perceptual preferences, independent of whether conceptualized in auditory or articulatory theories of speech perception, influence the kind of assimilation rules that languages of the world apply. As a consequence, compensation for assimilation occurs largely prelexically and is at that level independent of specific language experience.

Acknowledgments

This work was supported by a grant from the Dutch Scientific Organization (NWO, Project 048.011.046) in cooperation with the Hungarian Scientific Research Fund (OTKA, project number N 37282) to Leo Blomert. We thank Petra Vlamings and Dénes Szúcs for their help in running the experiment, and Wouter Jansen and Stefan Benus for answering questions about the phonetics of Hungarian liquid assimilation.

References

- Beddor, P. S., & Krakow, R. A. (1999). Perception of coarticulatory nasalization by speakers of English and Thai: Evidence for partial compensation. *Journal of the Acoustical Society of America*, 106, 2868–2887.
- Boersma, P. (1998). *Functional Phonology. Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.
- Boersma, P., & Weenink, D. (2001). Praat (Version4.0) [Computer software and manual]. Retrieved from <http://www.fon.hum.uva.nl/praat>
- Booij, G. (1995). *The phonology of Dutch*. Oxford, England: Clarendon.
- Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound*. Cambridge, MA: MIT Press.
- Clark, A. (1997). *Being there: Putting brain, body, and world together again*. Cambridge, MA: MIT Press.
- Coady J. A., Kluender, K. R., & Rhode, W. S. (2003). Effects of contrast between onsets of speech and other complex spectra. *Journal of the Acoustical Society of America*, 114, 2225–2235.
- Coenen, E., Zwitserlood, P., & Bölte, J. (2001). Consequences of assimilation for word recognition and lexical representation. *Language and Cognitive Processes*, 15, 535–564.
- Cutting, J. E., & Rosner, B. S. (1974). Categories and boundaries in speech and music. *Perception and Psychophysics*, 16, 564–570.
- Darcy, I. (2002). Online processing of phonological variation in speech comprehension: The case of assimilation. In S. Hawkins & N. Nguyen (Eds.), *Temporal integration is the perception of speech* (p. 32). Baixas, France: International Speech Communication Association.
- Dawkins, R. (1989). *The selfish gene* (2nd ed.). Oxford, England: Oxford University Press.
- de Boer, B. (2000). Self-organisation in vowel systems. *Journal of Phonetics*, 28, 441–465.
- Delgutte, B. (1997). Auditory neural processing of speech. In W. J. Hardcastle & J. Laver (Eds.), *The handbook of phonetic sciences* (pp. 507–538). Oxford, England: Blackwell.
- Ernestus, M. (2000). *Voice assimilation and segment reduction in Dutch*. Unpublished doctoral dissertation, Netherlands Graduate School of Linguistics, Utrecht.
- Evans, B., & Iverson, P. (2004). Vowel normalization for accent: An investigation of best exemplar locations in northern and southern British English sentence. *Journal of the Acoustical Society of America*, 115, 352–361.
- Farnetani, E. (1997). Coarticulation and connected speech processes. In W. J. Hardcastle (Ed.), *Handbook of phonetic sciences* (pp. 371–403). Oxford, England: Blackwell.
- Fowler, C. A. (1996). Listeners do hear sounds, not tongues. *Journal of the Acoustical Society of America*, 99, 1730–1741.
- Fowler, C. A., Best, C. T., & McRoberts, G. W. (1990). Young infants' perception of liquid coarticulatory influences on following stop consonants. *Perception and Psychophysics*, 48, 559–570.
- Fowler, C. A., & Brown, J. M. (2000). Perceptual parsing of acoustic consequences of velum lowering from information for vowels. *Perception and Psychophysics*, 62, 21–32.
- Fowler, C. A., Brown, J. M., & Mann, V. A. (2000). Contrast effects do not underlie effects of preceding liquids on stop-consonant identification by humans. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 877–888.
- Ganong, W. F. (1980). Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 110–125.
- Gaskell, M. G. (2003). Modelling regressive and progressive effects of assimilation in speech perception. *Journal of Phonetics*, 31, 447–463.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1996). Phonological variation and inference in lexical access. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 144–158.
- Gaskell, M. G., & Marslen-Wilson, W. D. (2001). Lexical ambiguity resolution and spoken word recognition: Bridging the gap. *Journal of Memory and Language*, 44, 325–349.
- Gaskell, M. G., & Marslen-Wilson, W. D. (1998). Mechanisms of phonological inference in speech perception. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 380–396.
- Gow, D. W. (2001). Assimilation and anticipation in continuous spoken word recognition. *Journal of Memory and Language*, 45, 133–159.

- Gow, D. W. (2002). Does English coronal place assimilation create lexical ambiguity. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 163–179.
- Gow, D. W. (2003). Feature parsing: Feature cue mapping in spoken word recognition. *Perception and Psychophysics*, 65, 575–590.
- Gow, D. W., & Im, A. M. (2004). A cross-linguistic examination of assimilation context effects. *Journal of Memory and Language*, 51, 279–296.
- Gussenhoven, C., & Jacobs, H. (1998). *Understanding phonology*. London: Arnold.
- Holt, L. L., Lotto, A. J., & Kluender, K. R. (2001). Influence of fundamental frequency on stop-consonant voicing perception: A case of learned covariation or auditory enhancement? *Journal of the Acoustical Society of America*, 109, 764–774.
- Howell, P., & Rosen, S. (1984). Natural auditory sensitivities as universal determiners of phonemic contrast. *Linguistics*, 21, 205–235.
- Hume, E., & Johnson, K. (2001). The interplay of perception and phonology. In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 3–26). New York: Academic.
- Hura, S. L., Lindblom, B., & Diehl, R. (1992). On the role of perception in shaping phonological assimilation rules. *Language and Speech*, 35, 59–72.
- Jakobson, R., Fant, G., & Halle, M. (1963). *Preliminaries to speech analysis*. Cambridge, MA: MIT Press.
- Jufrasky, D., Bell, A., & Girand, C. (2002). The role of the lemma in form variation. In C. Gussenhoven and N. Warner (Eds.), *Laboratory phonology 7* (pp. 3–34). Berlin, Germany: Mouton.
- Kohler, K. J. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 69–92). Dordrecht, The Netherlands: Kluwer.
- Kuhl, P. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, 50, 524–536.
- Ladefoged, P. (1996). *Elements of acoustic phonetics* (2nd ed.). Chicago: University of Chicago Press.
- Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *Journal of the Acoustical Society of America*, 29, 98–104.
- Ladefoged, P., & Maddieson, I. (1996). *Sounds of the world's languages*. Oxford, England: Blackwell.
- Lahiri, A., & Marslen-Wilson, W. (1991). The mental representation of lexical form: A phonological approach to the lexicon. *Cognition*, 38, 245–294.
- Lahiri, A., & Reetz, H. (2002). Underspecified recognition. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology 7* (pp. 637–676). Berlin, Germany: Mouton.
- Lerner, J., & Loewe, F. (1956). *My fair lady: A musical play in two acts*. New York: Coward-McCann.
- Liberman, A. M. (1996). *Speech: A special code*. Cambridge, Mass: MIT Press.
- Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling* (pp. 403–439). Dordrecht, The Netherlands: Kluwer.
- Lindblom, B. (1996). Role of articulation in speech perception: Clues from production. *Journal of the Acoustic Society of America*, 99, 1683–1692.
- Lotto, A. J., & Kluender, K. R. (1998). General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification. *Perception and Psychophysics*, 60, 602–619.
- Lotto, A. J., Kluender, K. R., & Holt, L. L. (1997). Perceptual compensation for coarticulation by Japanese quail (*Coturnix coturnix japonica*). *Journal of the Acoustical Society of America*, 102, 1134–1140.
- Magnuson, J., McMurray, B., Tanenhaus, M., & Aslin, R. N. (2003). Lexical effects on compensation for coarticulation: A tale of two systems? *Cognitive Science*, 27, 801–805.
- Mann, V. A. (1980). Influence of preceding liquid on stop-consonant perception. *Perception and Psychophysics*, 28, 407–412.
- Mann, V. A. (1986). Distinguishing universal and language-dependent levels of speech perception: Evidence from Japanese listeners perception of English ‘l’ and ‘r’. *Cognition*, 24, 169–196.
- Mann, V. A., & Repp, B. H. (1981). Influence of preceding fricative on stop-consonant perception. *Journal of the Acoustical Society of America*, 69, 548–558.
- Marslen-Wilson, W. D., & Warren, P. (1994). Levels of perceptual representations and processes in lexical access: Words, phonemes, and features. *Psychological Review*, 101, 653–675.

- Massaro, D. (1998). *Perceiving talking faces: From speech perception to a behavioral principle*. Cambridge, MA: MIT Press.
- McQueen, J. M., & Cutler, A. (2001). Spoken word access processes: An introduction. *Language and Cognitive Processes*, 16, 469–490.
- McQueen, J. M., Dahan, D. L., & Cutler, A. (2003). Continuity and gradedness in speech processing. In N. O. Schiller & A. S. Meyer (Eds.), *Phonetic and phonology in language comprehension and production: Differences and similarities* (pp. 39–78). Berlin, Germany: Mouton.
- Mitterer, H., & Blomert, L. (2003). Coping with phonological assimilation in speech perception: Evidence for early compensation. *Perception and Psychophysics*, 65, 956–969.
- Mitterer, H., Csépe, V., & Blomert, L. (in press). The role of perceptual integration in the perception of assimilated word forms. *Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*.
- Mitterer, H., & Ernestus, M. (2006). Listeners hear /t/s that speakers reduce. *Journal of Phonetics*, 34, 73–103.
- Näätänen, R. (1992). *Attention and brain function*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Näätänen, R. (1995). The mismatch negativity: A powerful tool for cognitive neuroscience. *Ear and Hearing*, 16, 6–18.
- Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, 38, 1–21.
- Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., et al. (1997). Language specific phoneme representation revealed by electric and magnetic brain responses. *Nature*, 385, 432–434.
- Näätänen, R., Schröger, E., Karakas, S., Tervaniemi, M., & Paavilainen, P. (1993). Development of a memory trace for a complex sound in the human brain. *Neuroreport*, 4, 503–506.
- Näätänen, R., Tervaniemi, M., Sussman, E., Paavilainen, P., & Winkler, I. (2001). “Primitive intelligence” in the auditory cortex. *Trends in Neuroscience*, 24, 283–289.
- Näätänen, R., & Winkler, I. (1999). The concept of auditory stimulus representation in cognitive neuroscience. *Psychological Bulletin*, 125, 826–859.
- Nearey, T. D. (1989). Static, dynamic, and relational properties in vowel perception. *Journal of the Acoustical Society of America*, 85, 2088–2113.
- Nolan, F. (1992). The descriptive role of segments: Evidence from assimilations. In G. J. Docherty & R. D. Ladd (Eds.), *Papers in Laboratory Phonology: II. Gestures segment, prosody* (pp. 261–280). Cambridge, England: Cambridge University Press.
- Ohalo, J. J. 1990. The phonetics and phonology of aspects of assimilation. In J. Kingston & M. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech* (pp. 258–275). Cambridge, England: Cambridge University Press.
- Olsson, M. (1992). *Hungarian phonology and morphology*. Lund, Sweden: Lund University Press.
- Otake, T., Yoneyama, K., Cutler, A., & Van der Lugt, A. H. (1996). The representation of Japanese moraic nasals. *Journal of the Acoustical Society of America*, 100, 3831–3842.
- Pastore, R. E. (1987). Categorical perception: Some psychophysical models. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 29–53). Cambridge, MA: Cambridge University Press.
- Pepperkamp, S. (2003). Phonological acquisition: Recent attainments and new challenges. *Language and Speech*, 46, 87–114.
- Phillips, C. (2001). Levels of representation in the electrophysiology of speech perception. *Cognitive Science*, 25, 711–731.
- Schröger, E. (1998). Measurement and interpretation of the mismatch negativity. *Behavior Research Methods, Instruments and Computers*, 30, 131–145.
- Schwartz, J.-L., Boë, L. J., Vallée, N., & Abry, C. (1997). Major trends in vowel system inventories. *Journal of Phonetics*, 25, 233–253.
- Sinnott, J. M., & Saporita, T. A. (2000). Differences in American English, Spanish, and monkey perception of the say–stay trading relation. *Perception and Psychophysics*, 62, 1312–1319.
- Siptár, P., & Törkenczy, M. (2000). *The phonology of Hungarian*. Oxford, England: Oxford University Press.
- Steriade, D. (2001). Directional asymmetries in place assimilation: A perceptual account. In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 219–250). New York: Academic.

- Sussman, H. M., Fruchter, D., Hilbert, J., & Sirosh, J. (1998). Linear correlates in the speech signal: The orderly output constraint. *Behavioral and Brain Sciences*, 21, 241–299.
- Traunmüller, H., & Eriksson A. (2000). Acoustic effects of variation in vocal effort by men, women, and children. *Journal of the Acoustical Society of America*, 107, 3438–3451.
- van Bezooijen, R. (1995). Sociocultural aspects of pitch differences between Japanese and Dutch women. *Language and Speech*, 38, 253–266.
- Weber, A. (2001). Help or hindrance: How violation of different assimilation rules affects spoken-language processing. *Language and Speech*, 44, 95–118.
- Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csepe, V., et al. (1999). Preattentive detection of vowel contrasts uses both phonetic and auditory memory representations. *Cognitive Brain Research*, 7, 357–369.