

# Compensation for phonological assimilation in perception: Evidence from Hungarian liquid assimilation

Holger Mitterer<sup>†</sup>, Valéria Csépe,<sup>‡</sup> and Leo Blomert<sup>†</sup>

<sup>†</sup> Universiteit Maastricht, Faculty of Psychology, Dept. of Cognitive Neuroscience

<sup>‡</sup> Hungarian Academy of Sciences, Institute for Psychology, Dept. of Psychophysiology, Budapest

E-mail: h.mitterer@psychology.unimaas.nl, csepe@cogpsyphy.hu, l.blomert@psychology.unimaas.nl

## ABSTRACT

Phonological assimilations add to the invariance problem. It is generally accepted that assimilated word forms are only recognized as an instance of the intended word in phonological contexts in which assimilation is allowed. We examined the mechanisms that drive this compensation for assimilation by presenting Hungarian morphologically complex words and non-words to Hungarian and Dutch listeners in an identification task, a discrimination task and a passive-oddball task during which the ERPs were measured. All three measures showed evidence for compensation for assimilation. This is unaffected by the lexical status of the assimilated (non)word. Results were similar for Dutch and Hungarian listeners, although Dutch listeners were not familiar with the assimilation rule. ERPs results indicate that acoustic details influence the compensation process, and replicate the absence of a native-language influence. These results indicate that auditory processes contribute to compensation for assimilation.

## 1. INTRODUCTION

Phonological assimilations have lately been the focus of investigations in two different fields. Linguists [1,2,3,4] try to elucidate the pattern of assimilation rules in the world's languages. From this perspective, it has been argued that assimilation rules are shaped by perceptual constraints. Hence, assimilation rules are the explicandum and 'perception' is the explicans. This contrasts with a psycholinguistic approach [5,6], which investigates how assimilated word forms are recognized. That is, how are listeners able to recognize, for instance, the word 'lean' if it has undergone assimilation as in the utterance 'leam bacon'? How do listeners compensate for the variance in the speech signal caused by phonological assimilations? In this word-recognition approach, what has to be explained is how the listener deals with assimilated forms in word recognition. In a way, this is an opposite approach to the linguistic one, as perception becomes the explicandum, while assimilation rules are taken as a given.

Not surprisingly, these two approaches have led to quite different proposals of how assimilated words are recognized. The psycholinguistic approach assumes that listeners implicitly learn the phonological assimilation

rules of their native language. This phonological knowledge allows them to recognize assimilated word forms. The learning algorithm takes into account the context-sensitivity of assimilation rules. That is, an assimilated form as 'leam' is accepted as an instance of 'lean' only in a context that allows the assimilation (i.e., 'leam bacon') but not in a context that does not allow this specific assimilation ('leam gammon'). However, this context effect arises as the listeners 'learn' the phonological assimilation rules of their native language.

In contrast, the linguistic/phonetic approach predicts that assimilated forms can be recognized without a need for prior exposure to the assimilation rule. It is, because it is generally difficult for the blue-print listener to make a difference between /n/ and /m/ in 'leam bacon', that the assimilation is allowed. Recently, two approaches have been proposed from a spoken-word recognition perspective, which are more in accordance with the linguistic/phonetic approach. Gow [7,8] proposed a feature-parsing account: According to this view, there is a general process of phonological feature-parsing. Feature parsing de-convolutes the overlapping activations of phonological features by means of general grouping principles [9]. Given that assimilated segments usually contain cues for both the original and the assimilated place of articulation, the assumption is that the evidence for the assimilated place of articulation is grouped together with the assimilating segment. In case of the utterance "lea<sup>n</sup><sub>m</sub> bacon", the evidence for the labial place of articulation of the word-final nasal in "lea<sup>n</sup><sub>m</sub>" is grouped with the assimilating labial segment /b/ in "bacon". This leaves only the 'correct' alveolar place of articulation associated with the word-final nasal in "lea<sup>n</sup><sub>m</sub>". Most importantly, the grouping mechanisms are supposed to be independent of language experience and it is conceivable that phonological systems may be constructed in a way so that they simplify the feature-parsing process [10]

Mitterer and Blomert [11] proposed an auditory locus of the 'compensation-for-assimilation processes.' This proposal is based on the framework of Kingston and Macmillan [12], in which the phonological categorization of a speech sound is achieved in two steps. First, basic auditory processes yield a representation of the input in a multi-dimensional perceptual space. Decision rules then associate regions of perceptual space with phonological categories. Context effects, as in the perception of assimilated utterances can

arise at both processing levels. First, it is possible that the context influences the early auditory processing of a given segment. This leads to different representation of the segment in the auditory perceptual space, which, in turn, may lead to different phonological categorization by a context-insensitive decision rule. In contrast, it is also possible that the early processing is context-insensitive, but then the decision rule is altered by context information. In this case, the representation of a given segment is the same in auditory perceptual space; however, the phonological categorization is context-dependent. It is argued that the context effect arises as a consequence of a perceptual integration of assimilated and assimilating segment on auditory levels of processing. This perceptual-integration account builds explicitly on the assumption that assimilation may be shaped by ‘perception.’ This in turn, allows basic auditory processes to account for what seems like a specific problem in spoken-word recognition, the recognition of assimilation words.

These three proposals, the phonological-inference account, the feature-parsing account, and the perceptual-integration account differ on three counts. The foremost question is at which level of processing compensation occurs. Both the phonological-inference and the feature-parsing account assume that compensation for assimilation occurs at a speech-specific processing stage at which the auditory input has been transformed into phonological features. However, the perceptual-integration account assumes an auditory locus of the effect.

A second question—related to the question of the processing level at which compensation for assimilation occurs—is whether compensation for phonological assimilation is influenced by the lexicon. It has been argued that compensation is more likely if it yields a word [6]. Such a ‘word-superiority effect’ is more likely if compensation occurs at a higher, language-specific processing level [13].

Finally, it is under debate whether compensation for a certain assimilation rule depends on experience with that rule. Clearly, a perceptual-integration account predicts that compensation for assimilation is independent from language experience. Assimilations are assumed to be constrained by general perceptual preferences independent of language experience. Similarly, the feature-parsing account assumes that learning-independent Gestalt principle account for compensation for assimilation. In contrast, phonological inference is assumed to be learned [6].

## 2. THE EXPERIMENTS

In order to investigate these issues we used multiple methods in a cross-linguistic design. A cross-linguistic design allows us to evaluate the impact of phonological knowledge on the perception of phonological assimilations. This is best achieved if quite different languages with different assimilation rules are investigated. Therefore, we investigated the perception of Hungarian utterances in

which a rule of liquid assimilation has been applied [14,15]. This assimilation rule allows a morpheme/word-final apical lateral to become an apical trill, if the next segment is an apical trill, too. Hence the Hungarian word ‘bal’ (Engl. ‘left’) may be pronounced as ‘bar’ in [ba<sup>l</sup>,ro:l] (Engl. from the left), but not in \*[ba<sup>l</sup>,na:l]. Given that Hungarian is a Fin-Ugric language, and the assimilation rule is different from the place assimilation rules in Germanic languages, probing perception of such assimilated utterances allows us to test the generality of the ‘compensation-for-assimilation’ phenomenon. In addition, we tested the perception of these utterances by Dutch listeners. The Dutch language is only very remotely related to Hungarian and knows no rule of liquid assimilation [16].

We [17,18] probed the perception of these utterances with three methods. First, we made use of a perceptual identification task, in which we expected that listeners will recognize the Hungarian word ‘bal’ (Engl. ‘left’) in the viably assimilated utterance [barro:l], adapted from ‘balrol’ (Engl. ‘from the left’). However, the [bar] in \*[barnal], an unviably assimilated variant of ‘balnal’ (Engl. ‘at the left’), is expected to be perceived as a deviation of the canonical ‘bal’. In order to test this, we used a natural utterance of ‘bal,’ created a continuum of stimuli ranging from ‘bal’ to ‘bar’ and concatenated these with the case suffix [nal] (Engl. ‘at the...’) or with the case suffix [ro:l] (Engl. ‘from the...’). These stimuli were presented to Hungarian listeners, and they were asked to identify the first syllable as either ‘bal’ or ‘bar’. Results show evidence of context-sensitive compensation for assimilation. Stimuli resembling [bar] were often identified as ‘bal’ in [barro:l], but as ‘bar’ in [barnal]. This replicated earlier results obtained with English [5,6,7,8], Dutch [11], and German [19] stimuli. Similar as the results for the Hungarian word [bal], Hungarian listeners often identified the nonword stimulus [zarro:l] as [zalro:l], while the stimulus \*[zarna:l] was nearly never identified as [zalna:l]. The results of Dutch listeners were similar, yet subtly different: Similar to Hungarian listeners, Dutch listeners had a steeper identification function when identifying the words in the unviable context [ba?nal] than in the viable context [ba?ro:l]. Unlike Hungarian listeners, they did not have a bias towards the canonical form. That is, while Hungarian listeners responded preferably with the canonical form [bal] or [zal] in the viable context, Dutch listeners were uncertain about how to identify the utterances in the viable context. They often identified the assimilated form [barro:l] as the canonical form [balro:l] and often identified the canonical form [balro:l] as the assimilated form [barro:l]

As a second method, we used a discrimination task to probe the perception of these Hungarian utterances. With this task, we hoped to constrain the level at which compensation for assimilation may occur. If compensation occurs at a language-specific level of processing, and auditory processing of assimilated utterances is independent of context, then there should be no effect of context on the performance in a discrimination task. If, however, compensation occurs at an auditory level, discrimination performance should be influenced by context. These

predictions, however, presume that discrimination performance is not influenced by phonological categorization of the stimuli. This seems to be the case if a four-interval oddity (4I-oddity) task is used [20]. With this task, between-category discrimination is not better than within-category discrimination. Results obtained with a 4I-oddity task revealed that discrimination was (much) more difficult in the viable context than in the unviable context [17]. This effect was qualitatively and quantitatively the same with Hungarian listeners discriminating Hungarian word, Hungarian listeners discriminating Hungarian nonwords, and Dutch listeners discriminating Hungarian words, which are Dutch nonwords.

Finally, we used a passive-listening task, in which listeners were exposed to oddball series while watching a silent movie [18]. Two oddball series were presented. In one series the frequent stimulus (85%) was [balro:l] and the deviant (15%) was the viable alternative pronunciation [barro:l], in the second series the frequent stimulus was [balnal] and the deviant stimulus was the unviable alternative pronunciation \*[barnal]. Simultaneously with the presentation of the oddball series, the electroencephalogram was measured in order to derive event-related potentials (ERPs). With this set-up, ERPs should reveal a MisMatch Negativity (MMN) for the deviant stimulus [21]. A context-sensitive compensation for assimilation should, in addition, yield a smaller MMN for the viable alternative pronunciation than for the unviable alternative pronunciation. The results, however, only showed such a context-sensitive MMN if the deviant stimulus did not contain a good example of an apical. That is, the good examples of apical trills used as deviants induced similar MMN in both contexts, viable and unviable. However, a context-sensitive MMN was observed if the deviant contained a non-prototypical example of an apical trill. Notably, this result was obtained with both, Dutch and Hungarian, listeners, and was independent of the lexical status of the assimilated form.

### 3. DISCUSSION

We investigated the perception of assimilated word forms with a special emphasis on a possible contribution of lexical processing and experience with assimilation rules. First of all, our results broaden the scope of the phenomenon ‘compensation for assimilation.’ Previously, ‘compensation for assimilation’ has only been shown for place assimilation rules in Germanic languages. The current results indicate a similar process in the perception of a liquid assimilation in a Fin-Ugric language. In addition, the ERP results confirm that compensation is an early perceptual process, occurring as the utterance is being heard.

The identification data show that the perception of assimilated utterances is independent from the lexical status of the assimilated form. In addition, the perception of assimilated utterances is largely independent of experience with an assimilation rule. However, there seems to be a

small effect of language experience. Native listeners are more likely to perceive the ‘canonical forms’ in the viable context. This may indicate that listeners learn which phonemes of their native language are prone to assimilate. When in doubt about the phonological class of a phonological segment, native listeners are more likely to perceive the phoneme, which may have undergone assimilation.

Secondly, we investigated the perception of assimilated utterances by means of a discrimination task. If compensation occurs at a language-specific processing level, there should be no effect of context on discrimination performance. This was, however, the case. Discrimination between [bal] and [bar] was more difficult in the viable context than in the unviable context. This result was also unaffected by, first, the native language of the listener, and second, whether the assimilated form was a word or not. Such a result supports the assumption that basic auditory mechanisms contribute to compensation for assimilation. In addition it buttresses, the assumption that assimilation rules may be shaped by general perceptual constraints [1,2,3,4].

The MMN results, in addition, highlight the role of acoustic/phonetic detail in the perception of assimilated utterances. Evidence for a context-sensitive perception of assimilated utterances was only obtained when the assimilated segment was not a prototypical example of the phonological class. This replicated earlier reports, which also indicated a critical role for acoustic/phonetic detail in the perception of assimilated utterances [7,8]. Above all, the results obtained with the MMN validate the conclusion that compensation for assimilation is a perceptual process, occurring during the presentation of the word. The MMN is in effect measured in the same time frame, 350-450 ms after word onset, as the case suffix is presented.

These results allow us to evaluate the three accounts for the perception of assimilated utterances. First of all, the perceptual-integration account turns out to be well supported. The context effects in the perception of assimilated utterances extend to a discrimination task, and are independent of specific language experience when measured with a MMN and a discrimination task. The results were also to a large degree independent of specific language experience when measured with an identification task. The only aspect of the current data that is difficult to reconcile with a perceptual-integration account is the effect that specific language experience had on the identification performance.

The feature-parsing account can account for the fact that most effects were independent of specific language experience. However, this account is difficult to reconcile with two aspects of our data. First of all, there is an effect of context on a discrimination task, indicating that context effects are located at the auditory, and not at the phonological processing level. In addition, it is difficult to explain in terms of feature-parsing why there is an effect of language experience on the perception of assimilated word forms.

Only this latter aspect of the current data set is in line with the predictions of a phonological-inference account. Given that there is an effect of specific language experience only in the identification, but not in the discrimination task, indicates that language experience modifies the processing of assimilated forms on the phonological, and not on the auditory level. However, it is difficult to reconcile the phonological-inference account with the findings on perception of assimilated forms that is the percept is to a large extent independent of specific language experience and the context-viability effects do not dilute in a discrimination task.

### 3. CONCLUSIONS

It seems that a process of perceptual integration of target and context drives the perception of assimilated utterances on auditory levels of processing. This basic processing is moderated by language experience, probably due to the acquisition of (implicit) phonological knowledge about assimilation rules of one's native language.

### REFERENCES

- [1] D. Steriade, "Directional asymmetries in place assimilation: A perceptual account", In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp.219-250), San Diego, CA: Academic Press, 2001.
- [2] K. Kohler, "Segmental reduction in connected speech in German: Phonological facts and phonetic explanations", In W.J.Hardcastle and A. Marchal (Eds.), *Speech Production and Speech Modelling* (pp. 69-92). Dordrecht: Kluwer, 1990.
- [3] J. J. Ohala, "The phonetics and phonology of aspects of assimilation", In J. Kingston and M. Beckman (Eds.), *Papers in Laboratory Phonology I: Between the grammar and the physics of speech* (pp. 258-275), Cambridge: Cambridge University Press, 1990.
- [4] S.L. Hura, B. Lindblom, and R. Diehl, "On the role of perception in shaping phonological assimilation rules", *Language & Speech*, **35**, pp. 59-72, 1992.
- [5] M.G. Gaskell and W.D. Marslen-Wilson, "Phonological variation and inference in lexical access", *Journal of Experimental Psychology: Human Perception and Performance*, **22**, pp.144-158, 1996.
- [6] M.G. Gaskell and W.D. Marslen-Wilson, "Mechanisms of phonological inference in speech perception", *Journal of Experimental Psychology: Human Perception and Performance*, **24**, pp.380-396, 1998.
- [7] D.W. Gow, "Does English coronal place assimilation create lexical ambiguity", *Journal of Experimental Psychology: Human Perception and Performance*, **28**, 163-179, 2002.
- [8] D.W. Gow, "Feature parsing: Feature cue mapping in spoken word recognition", *Perception & Psychophysics*, in press.
- [9] A. Bregman, "*Auditory Scene Analysis*", Cambridge, Mass.: MIT Press, 1990.
- [10] D.W. Gow and C. Zoll, "The role of feature parsing in speech processing and phonology", *MITWPL*, in press.
- [11] H. Mitterer and L. Blomert. "Coping with phonological assimilation in speech perception: Evidence for early compensation", *submitted*.
- [12] J. Kingston and N.A. Macmillan, "Integrality of nasalization and fl in vowels in isolation and before oral and nasal consonants - a detection-theoretic application of the Garner paradigm", *Journal of the Acoustical Society of America*, **97**, pp. 1261-1285, 1995.
- [13] M.A. Pitt and J.M. McQueen, "Is compensation for coarticulation mediated by the lexicon?", *Journal of Memory and Language*, **39**, pp. 347-370, 1998.
- [14] M. Olsson, "*Hungarian phonology and morphology*", Lund: Lund University Press, 1992.
- [15] P. Siptár and M. Törkenczy, "*The Phonology of Hungarian*", Oxford, UK: Oxford University Press, 2000.
- [16] G. Booij, G., "*The phonology of Dutch*", Oxford: Clarendon Press, 1995.
- [17] H. Mitterer, V. Csépe, and L. Blomert, "Perceptual integration drives the perception of assimilated word forms" *submitted*.
- [18] H. Mitterer, V. Csépe, and L. Blomert, "The perception of assimilated word forms as indexed by the MMN", *submitted*.
- [19] E. Coenen, P. Zwitserlood, J. Bölte, "Consequences of assimilation for word recognition and lexical representation", *Language and Cognitive Processes*, **15**, pp. 535 – 564, 2001.
- [20] B. Schouten, E. Gerrits, and A. van Hessen, "The end of categorical perception as we know it", *Speech Communication*, in press.
- [21] R. Näätänen and I. Winkler, "The concept of auditory stimulus representation in cognitive neuroscience", *Psychological Bulletin*, **125**, pp. 826-85, 1999.