Open your eyes and listen carefully. Auditory and audiovisual speech perception and the McGurk effect in aphasia

Klitsch, Julia Ulrike

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2008

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 30-09-2019
4 THE INFLUENCE OF PHONETIC FEATURES ON
AUDITORY WORD AND NON-WORD
DISCRIMINATION PERFORMANCE OF DUTCH
APHASIC SPEAKERS

4.1 Introduction
Phonetic features are acoustically characterized by the concurrence of particular cues. We
use the term ‘phonetic feature’ in a broad sense, without adhering to a particular
phonological theory. Psycho- and neurolinguistic studies, particularly for English, have
shown that these acoustic cues differ with respect to their perceptibility, at least under
certain circumstances. Auditory discrimination performance is influenced not only by
phonetic features, but also by the lexical status of the stimulus material (e.g. Blumstein,
Baker, & Goodglass, 1977). This experiment assesses how far the observed pattern
generalizes cross-linguistically; in this case the target language is Dutch. In the following
section, the acoustic cues that characterize the target phonetic features of this experiment
are summarized. Subsequently, some psycho- and neurolinguistic examples from the
literature are given. Finally, some examples from earlier research concerning the distinction
between words and non-words are described.

4.1.1 Phonetic features and their acoustic reality/properties in
speech perception
Phonetic features are not indivisible wholes, but they are each indicated by a number of
acoustic properties that are time-varying (not static), relative (not absolute), and graded
(not representing one particular value, but a continuum of values; Blumstein, 2004).

The three target phonetic features of this experiment are manner of articulation,
voicing, and place of articulation. They contrast within and between the three consonant
classes of plosives, fricatives, and nasals.
Manner of articulation distinguishes the three consonant classes. Manner of articulation is a coarse distinction that could be described by more fine-grained features, but this is unnecessary for the purposes of this experiment. Plosives are produced by blocking off the oral cavity as well as the nasal cavity (by raising the velum) for the air stream. Releasing the occlusion in the oral cavity results in the characteristic burst of sound when the air stream backed-up in the oral cavity can flow again.

Characteristic for fricatives is that they are produced by squeezing air through a narrow constriction at some point in the oral cavity, resulting in seemingly random, higher-pitched noise, differing in intensity and diffusion over frequency ranges.

Nasals are produced by blocking off the oral cavity for the air stream and lowering the velum so that air can escape through the nose. Since for nasals the air stream is never completely blocked off, as it is for plosives, nasals acoustically resemble sonorants and accordingly they are cross-linguistically nearly always voiced.

The acoustic cues to the feature voicing are numerous and context-dependent. Voicing distinguishes within the class of the obstruents, between voiced and voiceless plosives and fricatives, but it is not distinctive for nasals. As for nasals, air can stream unhampered through the nasal cavity, thus evading the application of the Aerodynamic Voicing Constraint (AVC). According to the AVC, voicing requires two basic characteristics: the vocal chords must be appropriately tense and adducted and air must flow through the vocal chords (Ohala, 1997). In the case of obstruents, on the other hand, an obstacle at some point in the oral cavity hampers the air stream, resulting in an accumulation of air in the oral cavity, increasing the air pressure. If the air pressure in the oral cavity approximates to the subglottal pressure, this condition infringes upon one basic requirement of the AVC – the airflow through the vocal chords. When airflow lapses below a certain level, voicing ceases because vocal chords stop vibrating. This is why per default voiceless obstruents are favored. Passive and active expansion of the vocal tract enables the AVC requirements to be met and voiced obstruents to be obtained. The further back the place of articulation of an obstruct, the less space there is for dispersion of accumulated air. Furthermore, voiced fricatives face another problem; voicing and frication are polar opposites concerning oral air pressure. While voicing requires that it is as low as possible, frication demands it to be as high as possible. Accordingly, voiced fricatives have much less fricative energy than voiceless ones (Pickett, 1980). Raphael (2005) tried to account for this fact by suggesting that during the production of a voiceless fricative the vocal folds are opened farther, so that more air can flow into the constriction in the oral cavity and, thus,
result in a more turbulent frication sound. For the voiced fricatives, however, the glottal airflow into the oral constriction is more limited, which reduces the frication turbulence in the oral cavity.

Studies with synthetic speech have shown that the VOT measure (Voice Onset Time) is the most basic acoustic cue to voicing in plosives. For example, Lisker and Abramson (1964) showed that in many languages, including Dutch, voice is present during the voiced plosive, whereas for voiceless plosives voicing only starts after burst release. But there are other attributes than VOT that contribute to the voice–voiceless distinction. For plosives it is widely agreed that the duration of the burst is the determinant for voicing. Slis and Cohen (1969) report that for Dutch, duration of the noise burst is 15ms longer for voiceless than for voiced plosives and its amplitude is up to 50% higher than for voiced plosives. Furthermore, the silent interval preceding the burst of voiceless plosives is 28 ms longer than for voiced plosives. Another cue to the voicing distinction are formant transitions at the moment when consonant and vowel meet: rapid transitions are indicative of voiceless plosives and slow transitions of voiced plosives (Delattre, 1962). For fricatives, the primary acoustic cue to voicing is phonation during frication and related to it is the intensity of frication; VOT seems less important.

**Place of articulation** in plosives has two important acoustic cues that are highly context-dependent. First, the burst frequencies associated with the stop release and, second, the formant transitions, especially that of F2 in proportion to the other formants (Ladeffoged, 2005; Raphael, 2005). For initial voiceless Dutch plosives, the burst is a much stronger cue for velar than for labial or alveolar consonants and its efficacy largely depends on the following vowel. For the voiced Dutch plosives, bursts were found to be stronger place cues for alveolar /d/ than for bilabial /b/. Again, the efficacy of the burst was largely influenced by the subsequent vowel. In general, bursts were found to be more effective cues to the place of articulation of voiceless than of voiced plosives (Smits, ten Bosch, & Collier, 1996).

The most salient cues to fricative place of articulation are frequency spectra and, to a lesser degree, formant transitions. More posterior fricatives are characterized by a concentration of energy in the higher frequency range. According to Raphael (2005) this higher frequency spectrum can be explained by the oral cavity that serves as a resonance cavity in front of the constriction. In the case of the anterior fricatives, this resonance cavity is missing, so that their intensities are relatively lower, resulting in a more evenly spread frequency spectrum. The other cue to fricative place of articulation (F2 transitions in the immediate vicinity of the frication) is disambiguating in those cases in which the spectrum
of fricative noise is ambiguous between phonemes (Harris, Hoffman, Liberman, Delattre, & Cooper, 1958).

The acoustic cue to the place of articulation of nasals is determined in the oral cavity; the location of the occlusion in the oral cavity determines the particular oral resonance, which is not a particularly strong indicator. Hence, formant transitions are the more salient cue to nasal place of articulation (e.g. Repp, 1986; Raphael, 2005).

To summarize, even though there are no invariant acoustic cues and although the known cues are context-dependent, the three features, manner of articulation, voicing, and place of articulation, are all characterized by particular acoustic cues or more or less complex combinations of different cues that indicate their occurrence. The next section describes psycho- and neurolinguistic studies that show how differently the acoustic cues contribute to auditory speech perception.

4.2 Phonetic features and their psycho- and neurolinguistic reality
Miller and Nicely (1955) analyzed perceptual confusion over English consonants in the presence of noise and low-pass filtering. They found that under these listening conditions, perception of the different features (they distinguished a.o. voicing and place of articulation) by non-brain-damaged speakers was affected to varying degrees. That is, the perception of, for example, voicing was not affected very much, whereas the correct perception of place of articulation was severely affected. These findings led Miller and Nicely to the conclusion that the perception of each individual feature is largely independent of the perception of other features, just as if each feature had its own perceptual channel.

Another study, also exploiting noisy listening conditions, was conducted by Sumby and Pollack (1954). They investigated how the availability of visual cues under noisy listening conditions can improve the auditory perception of different phonetic features. Using bisyllabic words embedded in white noise of differing degrees, they found that decreasing the speech-to-noise-ratio deteriorated the auditory identification abilities (25% correct). Providing a view of the speaker’s face articulating the same word clearly improved the identification performance (75% correct). At high speech-to-noise ratios, providing the visual cue had little influence; performance on auditory-only and audiovisual presentation was at ceiling.

The finding that visual cues can ease auditory speech perception when the auditory signal is ambiguous or degraded has also been confirmed in later studies (e.g. Summerfield,
1979; Grant, Ardell, Kuhl, & Sparks, 1985) investigating the underlying processes of spoken language perception and production. According to Summerfield (1987), the phonetic features manner of articulation and voicing are largely invisible, but additional visual information of the speaker’s vocal tract can contribute to the distinction or identification of minimal pairs that differ with respect to the place of articulation.

Gow and Caplan (1996) assessed auditory discrimination abilities in aphasic speakers, controlling for different contrasts. They distinguished between articulator-bound and articulator-free contrasts. The phonetic features voicing and place of articulation belong to the first group; manner of articulation, amongst others, belongs to the second group. The authors found a significant difference in auditorily discriminating articulator-bound and articulator-free contrasts, with the latter being better discriminated. With respect to the discrimination abilities of the two articulator-bound phonetic features, Gow and Caplan’s study was indeterminate. According to Gow and Caplan, past research on auditory phoneme discrimination in aphasic speakers mostly focused on the comparison between the phonetic features voicing and place of articulation. The results of these studies were conflicting; some researchers reported better discrimination abilities for place of articulation (e.g. Caplan & Aydelott-Utman, 1994), while others observed the opposite pattern – better discrimination abilities for the voicing feature (e.g. Blumstein et al., 1977).

Since temporal processing is crucial to consonant perception, and especially to the articulator-bound features, Gow and Caplan put forward the assumption that the aphasic deficit could basically consist of a deficit in integrating spectral information over time. Then place of articulation, being predominantly encoded by rapidly changing formant transitions, would pose a particular challenge to auditory speech perception, since it depends on the successful temporal integration of the incoming acoustic speech signal. The temporal deficit hypothesis put forward by Tallal and Newcombe (1978) represents the idea that speech comprehension deficits in aphasia are due to impairment in processing rapidly changing acoustic cues. The authors trained their participants with left-hemispheric brain damage due to missile-injury to identify contrasting pairs of synthetic three-formant syllables. The syllables contrasted with respect to second formant transition and were to be identified as either /ba/ or /da/. Furthermore, the pairs varied in transition length on all formants that were either short (40 ms) or long (80 ms). On the latter pairs more patients performed better than on the pairs with short transitions. The six patients showing difficulty with the pairs with the short transitions also displayed more difficulty on a sequencing task with non-linguistic stimuli, varying with respect to the interstimulus interval (ISI). Performance on this task had a high correlation with the results on the Token Test, which
led Tallal and Newcombe to conclude that impaired language comprehension results from a primary defect in temporal analysis.

However, Riedel and Studdert-Kennedy (1985) and Studdert-Kennedy (2002) cast doubt on Tallal and Newcombe’s hypothesis. In the first place, Riedel and Studdert-Kennedy challenge the interpretation that place of articulation judgments improve as a consequence of formant transition extensions. In this context they cite findings from research with non-brain-damaged listeners who perceived a shift in the manner of articulation (bilabial plosives became glides, corresponding as closely as possible with respect to place of articulation) rather than the place of articulation when formant transitions were extended from 30 ms to 60 ms (Liberman, Delattre, Gerstman, & Cooper, 1956; Miller & Liberman, 1979). To assess the question of whether it is actually the perceptibility of place of articulation that improves as a consequence of the extension of formant transitions, or whether an increase in performance is just the consequence of a shift to another phonetic contrast, Riedel and Studdert-Kennedy conducted an identification task and a discrimination task. They had twelve aphasic patients with left hemisphere CVA identify and discriminate paired CV-syllables. Stimulus pairs varied with respect to the duration of formant transition (short: 30 ms; long: 82 ms). Furthermore, there was one set of pairs with extended formant transitions, with the extension only affecting F2 and F3, carrying most information concerning the place of articulation, while F1 remained unchanged. The authors failed to replicate the results of Tallal and Newcombe. As with Blumstein, Tartter, Nigro, and Statlender (1984), the authors did not find any evidence that prolonging the formant transitions improves general aphasic performance. However, Blumstein and colleagues reported that the patient diagnosed with word deafness did profit from extended formant transition. Whereas the stimuli in the identification task (with a lengthened formant transition on all three formants that were closely matched to the stimuli in the Tallal and Newcombe study) did result in a variety of responses, the stimuli in which only the transitions of F2 and F3 were extended consistently generated /ba/ or /da/ responses.

Finally, Blumstein and colleagues (1977) and Riedel and Studdert-Kennedy (1985) reported that aphasic patients were better at discrimination than at identification. This is difficult to reconcile with the assumption that impairments in processing rapidly changing acoustic cues are the core of phonological performance impairment in aphasia.
4.3 Evidence from previous research for a distinction between words and non-words

Another issue of the current experiment concerns the influence of the lexical status of the stimulus material on auditory discrimination performance. Numerous studies suggest that words and non-words are processed differently, as indicated by the fact that in aphasia they can be independently impaired (e.g. deep dyslexia vs. surface dyslexia or selective impairments on non-word repetition). Blumstein and colleagues (1977) assessed the question of whether ‘complex discriminative hearing’ (Luria, 1970) can account for the auditory language comprehension deficit of Wernicke’s aphasics. They tested the auditory discrimination abilities of words and non-words of 25 aphasic patients. The results did not support Luria’s hypothesis, for two reasons. The first reason is that the patients with the most severe comprehension deficit were not the ones who were most severely impaired on the phoneme discrimination task. The other reason was that all patients were significantly more impaired on auditory non-word discrimination than on auditory word discrimination. If the auditory language comprehension deficit of Wernicke’s patients were attributable to a purely phonological disorder, then the lexical status of the stimulus pairs should not have had such an influence. The authors account for the differences in auditory discrimination performance of words and non-words by suggesting different processing strategies. In the case of words, decisions could be made in terms of meaning, whereas for auditory non-word discrimination decisions are purely based on the perceptual-phonological level.

A psycholinguistic study by Burton and Blumstein (1995) investigated the influence of stimulus naturalness and stimulus quality on the effect lexical status may have upon phonetic categorization. They found an overall lexical effect independent of stimulus naturalness, while stimulus quality seemed to contribute slightly to the occurrence of a lexical effect. That is, the lower the stimulus quality the more pronounced the lexical effect. In order to account for their findings the authors suggested that under ideal, no-noise listening conditions, reliance on acoustic information only is sufficient to accomplish task requirements. In conditions of noise (more comparable to the everyday listening condition), the listener may have to revert to supportive strategies – for example, lexical strategies – resulting in a more pronounced lexical effect. In line with this is the finding by Samuel (1981) who showed that the phoneme restoration effect (the auditory illusion of a phoneme that was actually deleted or replaced by noise in an auditory target) was strongly determined by the lexical status of this target; that is, the effect was much less pronounced in non-words. Behrman and Bub (1992) reported that in the domain of written language
processing, non-brain-damaged speakers (from now on nbd-speakers) showed a robust word superiority effect. Hence, different processing routes were assumed in the language processing models as, for example, suggested by Ellis and Young (1988).

The aim of the present experiment is to investigate how the three previously described phonetic features, *manner of articulation*, *voicing*, and *place of articulation*, influence the auditory discrimination performance of Dutch aphasic speakers. The question is whether the auditory perception of some phonetic features is more vulnerable in aphasia than the perception of others, and whether a ranking of perceptibility of these features can be established for Dutch. We expect the auditory discrimination performance of Dutch aphasic speakers to reflect differences in the perceptual vulnerability of the tested phonetic features. Furthermore, it is of interest to see if and how the Dutch data relate to the data of English-speaking aphasic patients (e.g. Gow & Caplan, 1996); that is, whether patterns generalize across languages.

Furthermore, it is assessed whether there is evidence for a lexical effect in the auditory phoneme discrimination of Dutch aphasic speakers. We expect to find evidence for such an effect and want to further inquire whether the lexical status of the stimuli is a general modifier of auditory discrimination performance or whether it interacts with some phonetic features but not with others.

4.4 Methods

4.4.1 Participants

All participants were native speakers of Dutch with a left hemisphere CVA at least three months prior to the experiment and diagnosed with aphasia. Furthermore, participants had to be right-handed prior to their CVA, hearing had to be normal, and they should not suffer from (severe) attention deficits. Speech therapists selected patients on the basis of the presence of a phonological output impairment as arising from the Dutch version of the AAT (a score of 4 or lower on the phonemic level of the AAT spontaneous speech rating scale), or as emergent during daily therapeutic intervention. After having given written consent, 13 aphasic patients participated in the experiment. Due to external factors that heavily influenced his performance, one patient had to be excluded from the analyses, leaving twelve patients, four of whom were female. Their mean age was 54.6 years (range: 33.9–72.8) and the mean time post onset was 10.8 months (range: 3–21 months). The individual
patients are described in Appendix A. According to the Dutch version of the Aachen Aphasia Test (AAT; Graetz, De Bleser, & Willmes, 1992), four of the patients were diagnosed with Wernicke’s aphasia, three were suffering from Broca’s aphasia, two patients were suffering from anomic and mixed aphasia, respectively, and one was classified as suffering from conduction aphasia.

Even though we used standardized test material, four additional nbd-native speakers of Dutch (mean age: 56.1 years) participated as a control group in order to ensure that any impairment on the tests is attributable to aphasia and not to the validity of the spoken digitized stimulus material.

4.4.2 Materials

The auditory word and non-word discrimination tests were taken from the Dutch version of the PALPA test battery (Bastiaanen, Bosje, & Visch-Brink, 1995) which is the only standardized test battery for Dutch. Both tests consisted of 72 CVC-stimulus pairs, half of which were identical, and the other half were minimal pairs that differed with respect to one phonetic feature. The difference was manifested by the contrast of either one of the three phonetic features, manner of articulation, voicing, and place of articulation (each feature: n = 12/test). Manner of articulation was investigated by contrasting /t/ and /s/, /k/ and /χ/, /r/ and /l/, /m/ and /b/, and /n/ and /l/. Voicing was assessed by the contrasts between /b/ and /p/, as well as /d/ and /l/. Since Dutch has final devoicing, the voicing contrast was only assessed in word-initial position. The place of articulation contrast was manifested in the opposition of /k/ and /p/, /k/ and /l/, /p/ and /l/, and /m/ and /n/. Manner and place of articulation contrasts were manipulated either in word-final position or in metatheses – for example, lor – rol (‘rag – role’). For each of the three contrast positions there were twelve instances per test.

The distribution of the contrast positions makes a straightforward comparison of word-initial and word-final contrasts difficult. This is why in the results section (see 4.5) the voicing feature, occurring in word-initial position only, will exclusively be compared to those instances of the features manner of articulation and place of articulation, respectively, that occurred in metatheses. Metatheses actually comprise a double contrast, occurring once in word-initial position and again in word-final position. The word-initial voicing contrasts will not be directly compared to the word-final occurrences of the features manner of articulation and place of articulation because a possible position effect may influence the result. This is because a word-initial contrast is usually assumed to be more easily
processed (for example to discriminate or identify) than a contrast in word-final position, irrespective of the feature contrast. At least for words occurring in connected speech, one reason could be that initial consonants are usually longer in duration and, hence, more salient than consonants occurring in word-final position (e.g. Lehiste, 1960; Quené, 1992).

The stimulus material was read out pairwise in a sound-proof booth by a male native speaker of Dutch. The stimuli were digitized at a sampling rate of 48kHz with 16-bit quantization. The recording volume was adjusted to 0dB (the reference level for the loudest sound to avoid clipping), with a mean value of -3.5dB. The frequency range was from 0 to 8 kHz with a volume of -∞ dB (= no sound) above 8 kHz.

4.4.3 Procedure and Scoring
Participants were seated at a table in a quiet room and asked to listen to the stimulus pairs. They were instructed to judge whether the stimulus pairs were the same or not by replying with ‘yes’ if the pairs were the same (‘Are the two stimuli the same?’) and with ‘no’ if they contrasted. In addition, it was explicitly stated in the instructions whether the stimuli were words or non-words. The order of presentation, auditory discrimination of words before auditory discrimination of non-words, or vice versa, was random. The stimulus-files were presented at a comfortable listening level from the hard disc of a PC-compatible laptop via loudspeakers. A trial consisted of one stimulus pair with an inter-stimulus interval (ISI) of 500 ms and an inter-trial interval (ITI) of 5000 ms. One repetition was allowed or a short break was inserted between trials, if requested by the participant.

There were three practice items that could be repeated if necessary. When the experimenter was sure the patient understood the test, the experimental stimuli were presented. Answers were noted by the experimenter on a protocol sheet and scored later. No more feedback was given once the trial started.

4.5 Results
Several analyses were conducted, the first with the participant group as a between-subject factor, the second with the within-subject factors ‘lexical status’ and ‘phonetic features’. The scores were used as dependent variables. All aphasic patients were pooled into one aphasic group since the type of aphasia did not have an influence on performance. ‘Phonetic feature’ was a three-level factor (manner of articulation, voicing, place of articulation), whereas ‘lexical status’ was a two-level factor (words, non-words).
Given the small number of participants non-parametric tests have been used for data analysis: the Mann-Whitney U test for between-group comparisons and Wilcoxon tests for within-group comparisons. The individual performance of the aphasic patients is given in Appendix B.

4.5.1 Aphasic vs. nbd-control group

In table 4.1, the group results are presented. A Mann-Whitney U test was run to compare the overall auditory discrimination performance of the aphasic and the nbd-control group. There was a significant difference in scores for the nbd-control group and the aphasic group \((z = -2.92, p = .004)\). This is the case for the auditory word discrimination task \((z = -2.96, p = .003)\) as well as for the auditory non-word discrimination task \((z = -2.92, p = .003)\).

Furthermore, the nbd-control group performed better on the contrasting stimulus pairs \((z = -2.92, p = .004)\), which is reflected by each of the phonetic features *manner of articulation* \((z = -2.08, p = .037)\), *voicing* \((z = -2.87, p = .004)\) and *place of articulation* \((z = -2.93, p = .003)\). Finally, performance of the nbd-control group was at ceiling, so no more comparisons were run within this group.

<table>
<thead>
<tr>
<th>Table 4.1</th>
<th>Performance comparisons between aphasic group and nbd-control group, (p &lt; .05^*)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stimulus category</strong></td>
<td><strong>Percentage correct</strong></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>99.65</td>
</tr>
<tr>
<td>Auditory word discrimination</td>
<td>100</td>
</tr>
<tr>
<td>Auditory non-word discrimination</td>
<td>99.65</td>
</tr>
<tr>
<td><em>Manner of articulation</em></td>
<td>100</td>
</tr>
<tr>
<td><em>Voicing</em></td>
<td>98.95</td>
</tr>
<tr>
<td><em>Place of articulation</em></td>
<td>98.95</td>
</tr>
</tbody>
</table>
4.5.1.1 Influence of lexical status and phonetic features on aphasic discrimination performance

The data on the influence of lexical status are given in table 4.2. Wilcoxon tests were run to evaluate the influence of lexical status on the auditory discrimination performance of the aphasic patients. The scores obtained on auditory word discrimination were significantly higher than those on auditory non-word discrimination ($z = -3.07$, $p = .002$). This was true for all three phonetic features, manner of articulation ($z = -2.23$, $p = .026$), voicing ($z = -2.97$, $p = .003$), and place of articulation ($z = -2.79$, $p = .005$).

This overall difference between words and non-words resulted from the contrasting minimal pairs because in the auditory non-word discrimination task, significantly more minimal contrasts were missed than in the auditory word discrimination task ($z = -3.06$, $p = .002$). The number of false alarms (rejections even though the stimuli were identical) did not differ between the two tasks in the no-contrast condition ($z = -5.4$, $p = .591$) in which the minimal pairs were identical.

In summary, the difference in performance on words and non-words was due to the fact that more contrasting minimal pairs were missed in the non-word discrimination task than in the word discrimination task, reflecting an interaction between the two factors of lexical status and phonetic feature.

| Table 4.2 Influence of lexical status on patient performance; comparison between performance on the auditory word and non-word discrimination tasks, $p < .05^*$ |
|-------------------------------------------------|-----------------|-----------------|
| Stimulus category                               | Percentages correct |
|                                                 | Auditory word discrimination task | Auditory non-word discrimination task |
| Overall phonetic features                       | 93.04            | 82.88*          |
| Manner of articulation                          | 93.07            | 88.19*          |
| Voicing ($n = 12$)                              | 90.29            | 70.82*          |
| Place of articulation ($n = 12$)                | 86.82            | 55.55*          |
| Contrasting condition (difference to 100% correct = misses) | 90.04            | 70.38*          |
| No-contrast condition (difference to 100% correct = false alarms) | 96.06            | 95.35          |
The influence of phonetic features on auditory word and non-word discrimination performance of Dutch aphasic speakers

Subsequently, the influence of phonetic features on the performance of the patients was analyzed (Table 4.3). Running a Wilcoxon-test to evaluate the respective influence of the phonetic features, manner of articulation and place of articulation, on aphasic performance showed that there was a significant difference in performance between the two features \(z = -2.94, p = .003\).

This difference is present in the auditory non-word discrimination task \(z = -3.08, p = .002\) but fades out in the auditory word discrimination task \(z = -1.15, p = .250\).

<table>
<thead>
<tr>
<th>Feature contrasts</th>
<th>Overall</th>
<th>Auditory word discrimination</th>
<th>Auditory non-word discrimination</th>
</tr>
</thead>
<tbody>
<tr>
<td>MoA – PoA</td>
<td>90.63 – 71.18*</td>
<td>93.07 – 86.82</td>
<td>88.19 – 55.55*</td>
</tr>
<tr>
<td>MoA&lt;sub&gt;meta&lt;/sub&gt; – Voi</td>
<td>95.38 – 80.56*</td>
<td>100 – 90.29*</td>
<td>93.05 – 70.82*</td>
</tr>
<tr>
<td>Voi - PoA&lt;sub&gt;meta&lt;/sub&gt;</td>
<td>80.56 – 84.43</td>
<td>90.29 – 92.6</td>
<td>70.82 – 72.21</td>
</tr>
</tbody>
</table>

As mentioned in the material section, the feature voicing was only compared to those instances of the features manner of articulation and place of articulation that appeared in metatheses.

Finally, the impact of the phonetic features manner of articulation and voicing on the outcome of the auditory discrimination tasks was analyzed (Mann Whitney U tests). There was a significant difference between auditory discrimination performances on the features manner of articulation and voicing \(z = -2.55, p = .011\). This finding held for both the auditory non-word discrimination task \(z = -2.93, p = .003\) and the auditory word discrimination task \(z = -2.26, p = .024\).

Comparing performance on the two features voicing and place of articulation by means of a Wilcoxon test did not reveal significant differences, neither overall \(z = -1.41,
p = .158) nor in either of the two tasks, auditory word (z = -.53, p = .595) or auditory non-word discrimination (z = -.46, p = .645).

To control for a possible effect of contrast position on the phonetic features place of articulation and manner of articulation we conducted a post-hoc analysis for the two positions of occurrence (Table 4.4) – word-final position and metathesis – with position being a two-level factor.

A Wilcoxon-test was conducted to investigate whether there was an effect of contrast position. For the feature place of articulation there was an overall effect of position with the contrasts in metatheses being significantly better detected than those in word-final position (z = -3.06, p = .002). The same influence of contrast position was evident, independent of the lexical status of the stimulus material (Auditory word discrimination: z = -2.59, p = .010; auditory non-word discrimination: z = -2.68, p = .007).

For the feature manner of articulation, statistical analysis revealed a significant overall position effect (z = -2.25, p = .024). The same could be observed in the auditory word discrimination task (z = -2.26, p = .024), whereas performance on the two contrast positions in the auditory non-word discrimination task just missed being significantly different (z = -1.81, p = .070).

<table>
<thead>
<tr>
<th>Table 4.4</th>
<th>Comparisons between the two contrast positions (word-final position and metatheses) for the features manner of articulation and place of articulation, p &lt; .05*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Percentages correct</td>
</tr>
<tr>
<td></td>
<td>Overall</td>
</tr>
<tr>
<td></td>
<td>Final</td>
</tr>
<tr>
<td>MoA</td>
<td>87.78</td>
</tr>
<tr>
<td>PoA</td>
<td>49.08</td>
</tr>
</tbody>
</table>

In summary, the aphasic group performed on all variables at least two standard deviations worse than the nbd-control group; that is to say, aphasic performance was impaired on all variables.

Whereas performance for the nbd-group was at ceiling on all variables, the aphasic group showed a significant effect of lexical status. Moreover, the auditory discrimination performance of the aphasic group was influenced by the phonetic features that manifested
the contrast. Auditory discrimination of manner of articulation seemed to be least impaired, followed by voicing and place of articulation. Furthermore, there is an effect of contrast position.

4.6 Discussion

First, the question concerning phonetic features and their influence on aphasic auditory discrimination performance is addressed. Subsequently, the influence of the lexical status is discussed.

The aim of this experiment was to assess whether in Dutch there are phonetic features that are more vulnerable to aphasic damage in auditory discrimination, as others have found for English. As with the English-speaking patients (cf. Gow & Caplan, 1996), the Dutch patients in this experiment were overall significantly better at discriminating manner of articulation than either voicing or place of articulation.

The comparison of the feature manner of articulation (when occurring in metatheses) and the feature voicing (in word-initial position) revealed an effect of phonetic feature contrast on auditory discrimination performance of the aphasic group, at least when non-word stimulus pairs were to be discriminated. This suggests that auditory discrimination of manner of articulation was superior to auditory discrimination of voicing.

Comparing performance on the contrast detection for the features voicing and place of articulation occurring in metatheses showed no reliable differences. Still, we want to tentatively suggest that auditory discrimination of the feature place of articulation was more vulnerable to aphasic impairment than was the feature voicing. Earlier studies revealed that aphasic patients’ speech perception was especially vulnerable if there was only a single feature contrast as opposed to a multiple feature contrast (e.g. Baker, Blumstein, & Goodglass, 1981; Blumstein, et al., 1977; Miceli, Caltagirone, Gainotti, & Payer-Rigo, 1978; Miceli, Gainotti, Caltagirone, & Masullo, 1980; Blumstein, 1994). Thus, in the current experiment it seems appealing to equate the recurring contrast of the features place of articulation and manner of articulation in metatheses to a multiple feature contrast. Whereas multiple feature contrasts represent a qualitative difference (e.g. in voicing and place of articulation: /d/ vs. /p/), the repetition of the same single feature contrast is considered to represent a quantitative difference (e.g. pijt vs tijp; the same contrast occurs twice, in word-initial and in word-final position). Yet, it seems plausible to assume a facilitative effect of the latter as well. The post-hoc finding, that auditory discrimination performance on the features place of articulation and manner of articulation was
significantly better when these features contrasted in metatheses than when the contrast was in word-final position, would be in support of this assumption.

The lack of a statistically significant difference between voicing and place of articulation might thus be a result of enhanced auditory discrimination performance on the latter feature, which benefited from its repetitive occurrence in metatheses. A direct comparison between place of articulation just in word-final position, and voicing, could not be conducted since it is confounded by contrast position (word-final vs. word-initial).

Without drawing too strong a conclusion, the data suggest that auditory discrimination of the feature place of articulation may indeed be more vulnerable to aphasic impairment than perception of the voicing feature. This is because auditory discrimination performance on the feature place of articulation in metatheses, despite its repetitive occurrence, does not significantly outplay performance on the feature voicing. Thus, the resulting ranking of perceptual distinctiveness of the three distinguished phonetic features for aphasic speakers of Dutch suggests that, as in English, manner of articulation is least impaired. As in English, auditory discrimination of the other two features, voicing and place of articulation, is more impaired. However, limits in the stimulus design prevented the satisfactory establishment of a distinct ranking between the latter two features. Yet, it is tentatively put forward here that in Dutch, the feature place of articulation is most detrimental to auditory discrimination performance in aphasia. This coincides with the findings of some studies on English (e.g. Blumstein et al., 1977), while it contradicts others (Caplan & Aydelott-Utman, 1994).

The data of this experiment do not allow for a definition on acoustic grounds of what caused the varying degrees of vulnerability to aphasic impairment of the three phonetic features in auditory discrimination. Thus, it remains unclear whether there were certain acoustic cues that, in general, were particularly difficult to perceive for the aphasic system. Alternatively, the possibly reduced perceptibility of certain acoustic cues could also be caused by phonetic context and, hence, would only occur in certain vowel contexts or word positions. It is also conceivable that the integration of different acoustic cues was evoking the perceptual impairment, especially for place of articulation, as this feature is particularly complex because it is characterized by the interplay of numerous context-dependent acoustic cues.

Before starting to speculate what might have been the cause for a specific acoustic cue to be the most vulnerable to impaired discrimination performance, especially place of articulation, the attention should be drawn to the other main finding of the current experiment. That is, the aphasic auditory discrimination performance, in general, improved
significantly when the target stimuli were lexical words of Dutch; that is, the patients showed a word superiority effect, as was previously reported in the literature (e.g. Blumstein, 1994; Blumstein et al., 1977). The improvement in auditory discrimination performance was especially reflected by the two features whose perception was more vulnerable overall to aphasic damage; these two features are voicing and place of articulation. Furthermore, there was a strong effect of contrast position, independent of the lexical status of the stimulus material. That is, for place of articulation as well as manner of articulation auditory discrimination performance was significantly better when either one of the two features contrasted in metatheses than in word-final position. Thus, the question arises as to how far the discrimination impairment was acoustic in nature in the first place.

If it were a pure acoustic-phonetic impairment, word and non-word stimuli should be affected alike, at least at the earliest stages of auditory speech processing. In addition, the word superiority effect casts doubt on the temporal-deficit-hypothesis by Tallal and Newcombe (1978). If speech comprehension difficulties in aphasia did indeed arise as a consequence of impairment in processing rapidly changing acoustic cues, this should be true for all stimulus pairs, irrespective of lexical status and contrast position. While for auditory non-word pairs a discrimination decision can only be made on phonological grounds, in the case of auditory word pairs the aphasic patients seemed to employ lexical strategies in order to make the discrimination decision (cf. Blumstein et al., 1977). This would account for the observed lexical effect and, thus, seems most compatible with an auditory speech-processing model that assumes (interactive) top-down processing. However, the position effect that was also evident with the auditory word pairs suggests that, in this case, a top-down flow of information from the higher lexical level may be supportive. Yet, this cannot be the only decisive factor in the discrimination process. This latter finding is more suggestive of the idea that phonological level, lexical level, and short-term memory capacities contribute to auditory discrimination.

The current finding, which suggests that the feature place of articulation was particularly detrimental to aphasic auditory discrimination performance, is in line with findings from psycholinguistic studies with nbd-speakers. That is, reduced auditory perception abilities, namely for the feature place of articulation, is exactly what was found for nbd-listeners under noisy listening conditions (Sumby & Pollack, 1954). The nbd-control group performed at ceiling on the stimulus material of the present experiment, suggesting that there was no noise inherent to the stimuli or the procedure. Rather, the lower auditory discrimination performance of the aphasic participants is attributed to the aphasic impairment. That is, it is suggested that the aphasia causes some kind of system-internal
noise. It is this ‘aphasic noise’ that is considered to cause the lower overall discrimination scores and that is particularly detrimental to the accurate perception of the feature *place of articulation*.

In previous studies with nbd-speakers it was found that visual cues from the speaker’s face can ease auditory speech perception when the auditory signal is noisy – that is, when it is ambiguous or degraded (Grant, Ardell, Kuhl, & Sparks, 1985; Miller & Nicely, 1955; Sumby & Pollack, 1954). On the contrary, the speaker’s face had little influence at high speech-to-noise ratios; under these conditions, the performance was similar on both auditory-only and audiovisual presentation. In noise, however, it is *place of articulation* that benefits most from visual cues from the speaker’s face (Summerfield, 1987).

As a matter of course, noisy listening conditions cannot be equated to the presence of aphasia (which is said to resemble some kind of ‘system-inherent noise’) and this is not what is claimed. Yet, it is very possible that the auditory speech perception of aphasic speakers does improve with audiovisual stimulus presentation. The greatest improvement, then, would be expected for the feature *place of articulation*, since the other two features, *manner of articulation* and *voicing*, are largely invisible. Providing the (aphasic) listener with visual cues from the speaker’s face can be seen as an alternative or complement to temporal and spectral stimulus manipulations of the speech signal to improve spoken speech perception in aphasia. The effect of visual articulatory cues from the speaker’s face on aphasic speech perception abilities will be assessed in the second experiment, a McGurk experiment that also assesses auditory, visual, and audiovisual identification performance.