1 General Introduction

The topic of this study is speech perception in three different populations of Dutch native speakers: an aphasic group, an age-matched group, and a student group. Speech perception is investigated by means of an auditory discrimination task (AX-paradigm) and an identification task in the auditory, visual, and audiovisual modalities. The audiovisual modality distinguishes between congruent and incongruent stimuli (McGurk stimuli). In addition, the stimulus material of both tasks is controlled for the variable lexical status.

Auditory discrimination tasks conducted in English uniformly showed that aphasic speakers are better at the auditory discrimination of manner of articulation than at the discrimination of either voicing or place of articulation (e.g. Gow & Caplan, 1996). However, the ranking of the two features that are more difficult to discriminate for aphasic speakers of English was less clear (cf. Blumstein, Baker, & Goodglass, 1977; Caplan & Aydelott-Utman, 1994). In studies with non-brain-damaged speakers (nbd-speakers), noise has been shown to be detrimental to speech perception in general and especially to the perception of the feature place of articulation (e.g. Miller & Nicely, 1955). Another finding was that the lower the speech-to-noise ratio, the greater the influence of visual speech information (Sumby & Pollack, 1954). That is, decreasing the speech-to-noise ratio was shown to reduce auditory speech perception while at the same time the influence of visual speech information increased.

The current study aims to discover the pattern of the speech perception impairment of aphasic speakers of Dutch. A number of variables that have been reported to influence speech perception are controlled to allow for the best possible description and to enable cross-linguistic comparison. Identifying the variables that are impaired in aphasic speech perception is one aspect, but of even more clinical relevance is the way in which the observed speech perception impairment can be mitigated. Different endeavors have been made to increase aphasic speech perception, without achieving sustained success. Examples are manipulations of the spectral and/or temporal cues in the speech signal (e.g. Tallal & Newcombe, 1978; Blumstein, Tartter, Nigro, and Statlender, 1984) or the recommendation to speak more slowly and in a silent surrounding. Yet, as has been shown with nbd-speakers, visual speech information has been proven to make a natural contribution to improve speech perception in noisy listening conditions. Although the experiments in this study are administered in a silent environment, it is investigated if speech perception of aphasic patients can be improved by providing visual speech information. If this is the case,
relatives and attendants of aphasic speakers should (in addition) be instructed to articulate clearly (not more loudly) while looking at the conversational partner with aphasia to provide him/her with as many speech cues as possible.

This study assesses these questions with two experiments that are described in two separate chapters: chapter 4 and chapter 6. Prior to the experimental chapters, a theoretical background is sketched, with chapter 2 giving an overview of auditory and audiovisual speech perception in non-brain-damaged speakers and speakers with aphasia. In chapter 3, models of auditory and audiovisual speech perception are briefly outlined and evidence from neuroimaging studies for one or the other potential models is described. Furthermore, a study by Boersma (2006) is illustrated that tries to account for the McGurk effect in terms of Optimality Theory (OT). Chapter 4 describes the first experiment in which the auditory discrimination abilities of twelve aphasic speakers of Dutch are investigated with word and non-word minimal pairs. The dissimilar pairs differed minimally with respect to one of the three phonetic features—manner of articulation (MoA), place of articulation (PoA), or voicing (Voi). Position of contrast was controlled and occurred either word-initially (Voi only), word-finally (MoA and PoA), or in metatheses (MoA and PoA). The fifth chapter introduces the McGurk effect and provides information relevant for the second experiment that is based on the results of the first experiment. Chapter 6 is concerned with the McGurk experiment, an identification task that distinguished between four conditions, the auditory-only condition (AO), the visual-only condition (VO), the audiovisually congruent condition (AV), and the audiovisually incongruent McGurk condition. The McGurk experiment made use of the three-alternatives-forced-choice paradigm (3 AFC). Furthermore, in each of the four testing conditions, the lexical status and phonetic context of the stimulus material were controlled. Finally, chapter 7 summarizes both experiments of the current study, discusses the main findings, and provides a proposal for further investigation.

Before turning to the next chapter, the remainder of this chapter introduces four topics that are relevant for the current study.

1.1 Cognition and aging
Whereas pathological aging (for example, with Alzheimer’s disease) is associated with severe neuroanatomical deterioration and the reduction of synaptic density, mild cognitive changes, as observed during normal aging, are attributed to a more moderate decline of cortical structures and neuronal connections (Schneider, Rowe, Johnson, Holbrook, & Morrison, 1996). In addition, Li and Lindenberger (2006) give an overview of how
neuromodulation may be changing with increasing age. Neuromodulation describes the regulation of multiple (groups of) neurons that are widely spread across the nervous system by neurotransmitters (for example, dopamine or serotonin). This suggests that age-related cognitive changes can be caused by structural cortical changes and/or by changes in the (interactions of) various transmitter systems. Hence, the reduction of dopamine receptors has been suggested as a neurochemical correlate of unimpaired cognitive aging. Li and Lindenberger give several pieces of evidence from the literature that the decline of dopamine receptors starts at age 20 and that the deterioration process in different cortical regions continues with 7–11% per decade of adult life (Wong, Young, Wilson, Meltzer, & Gjedde, 1997). Experimental studies with humans suggest a functional relationship between a reduction in dopamine receptors and a decline in cognitive abilities (for example, processing speed or episodic memory: Bäckman, Ginovart, Dixon, et al., 2000). In summary, Li and Lindenberger describe the possible consequences of an age-related decrease in dopaminergic modulation which negatively influences the efficiency of widely distributed cortical neuron activity and, hence, causes less distinct internal representations as elicited by different stimuli.

Attention is one of the many cognitive functions and processes that are affected by the age-related reduction of neuromodulation. Two types of attention may be of particular interest for the present study: selective attention and divided attention. Selective attention stands for the successful focus onto one relevant source of information, while present distractors are successfully ignored. Divided attention places an even higher demand on the cognitive system, since it denotes the ability to successfully divide or share the available processing resources between two or more competing perceptual tasks or cognitive processes (Sarter & Turchi, 2002). That attention decreases as age increases, has been reported in numerous studies (for example, Nyberg, Nilsson, Olofsson, & Bäckman, 1997, for the vulnerability of divided attention to age effects).

1.2 Aphasia

Aphasia is an acquired language disorder resulting from focal brain damage after completed language acquisition. Aphasia has, thus, to be differentiated from language acquisition disorders, language disorders resulting from hearing impairments, or language disorders related to psychological disorders. All aphasic participants in this study were aphasic due to vascular incidents; aphasia as a consequence of tumor, brain trauma, encephalitis, or brain atrophy was an exclusion criterion.
Aphasia can affect all linguistic levels – morphology, phonology/phonetics, semantics, syntax – in perception and production, and in spoken and written language alike. That is, aphasias are assumed to be supra-modal (Huber, Poeck, Weniger, & Willmes, 1983). The degree and severity of impairment, in general as well as at a specific level, may vary from patient to patient. Huber and colleagues (1983) distinguish four standard syndromes of aphasia: anomic aphasia, Wernicke’s aphasia, Broca’s aphasia, and global aphasia, each of which is associated with a specific linguistic symptom complex. Anomic aphasia is characterized by more or less severe word finding difficulties that are sometimes compensated for with lengthy circumscriptions. Grammatically the speech of individuals with anomic aphasia is intact, but the frequent circumscriptions make the speech appear somewhat empty. Auditory speech comprehension is (relatively) preserved. Wernicke’s aphasia is characterized by fluent, well-articulated, but often meaningless speech due to many phonological and/or semantic paraphasias and neologisms. Speech comprehension, oral and written, is severely impaired in individuals with Wernicke’s aphasia. The speech of individuals with Broca’s aphasia is slow, nonfluent, and often agrammatic with relatively better-preserved auditory speech comprehension abilities. Global aphasia is the most severe form of aphasia. Individuals with global aphasia have no, or severely limited, speech which often comprises nothing but recurring utterances. Speech comprehension in these individuals is also severely impaired.

There are efforts to neuroanatomically relate the aphasic standard syndromes to a circumscribed area of brain damage. For example, in 1861 the French neurologist Paul Broca described a patient, TanTan, who could not utter anything except his name but who was quite well able to comprehend speech. In a post-mortem investigation, Broca stated that TanTan had suffered from brain damage in the neuroanatomical area known as the motor speech center, spanning pars opercularis and to some extent pars triangularis of the Gyrus frontalis inferior in the language-dominant hemisphere (Trepel, 1995). This region, in Brodman’s areas 44 and 45 (although the involvement of Brodman’s area 45 is not unanimously agreed upon), has since been known as Broca’s area. Accordingly, damage to this region in the language-dominant hemisphere is associated with symptoms characteristic for Broca’s aphasia. In addition to the linguistic impairment, individuals with Broca’s aphasia often have a hemiparesis of face and limbs in the contralateral side of the brain damage. Via the arcuate fasciculus, Broca’s area is connected to Wernicke’s area, the sensory speech center, in the secondary auditory cortex in the left temporal lobe, and spanning Brodman’s areas 42 and 22. In 1874, Carl Wernicke, a German neurologist, described a patient who had a lesion in this area and who showed the characteristics of
what has since been known as Wernicke’s aphasia. However, in cognitive neuropsychology the clear relationship between neuroanatomical localization of the lesion site and expected neurolinguistic patterns is contested due to the double dissociations of the lesion site and the observed language disorder.

1.3 Phonetic features

In this study, consonants are the critical phonemes that constitute the minimal contrasts in the auditory discrimination task and in the identification task. That is why the present work only discusses those phonetic features that are relevant for the distinction between consonants.

It is characteristic for consonants that during articulation the airflow is obstructed at some point. In articulatory phonetics, the distinction is made between active articulators, which move to cause the obstruction (usually the tongue), and passive articulators, which indicate the place in the articulation tract at which the obstruction is caused. The first phonetic feature constituting a minimal contrast in the tasks of this study is the feature place of articulation. This feature is characterized by the interaction of active and passive articulators. Three different places of articulation, that is, of obstruction are distinguished in this study: (bi)labial articulation, articulation at the alveoles, and articulation involving a constriction at the uvula.

In addition, the degree of obstruction characterizes whether the obstruction is complete, as for stop consonants or plosives, or partial, as for fricatives, for example. Furthermore, when the oral cavity is completely blocked for the air stream but air can escape via the nose, a nasal is being produced. The second phonetic feature central to this study is the feature manner of articulation to which these three consonant characteristics belong.

The third relevant phonetic feature of this study, voicing, concerns the vocal chords and whether they vibrate during articulation. When the vocal chords vibrate during the release of the obstruction (that is, when voice is present at the onset of the release: Voice Onset Time, VOT), a consonant is supposed to be voiced. With a delay between release of the occlusion and onset of voicing, that is, a longer VOT, a consonant is called ‘unvoiced’ or ‘voiceless’.
1.4 Frequency Formants

Formants reflect accumulations of acoustic energy in the speech signal and usually occur at intervals of ca. 1000 Hz. The formants of oral phonemes are numbered in ascending order. A(n oral) vowel generally consists of four to six frequency formants, but usually the first two formants, F1 and F2, sufficiently characterize a vowel to distinguish it from other vowels. F1 varies as a consequence of vertical tongue movement; the lower the tongue, the higher the F1 value. For male speakers of Northern Standard Dutch, the F1 values of vowels vary between ca. 250 Hz and ca. 700 Hz (Adank, Van Hout, & Smits, 2004). F2 reflects tongue movement in the horizontal dimension. The more the highest part of the tongue moves to the front during articulation, the higher the F2 value. F2 is further influenced by liprounding, with unrounded vowels causing lower F2 values. The range of F2 values for vowels articulated by a male speaker of Northern Standard Dutch varies between ca. 800 Hz and ca. 2000 Hz (Adank et al., 2004).

However, not only vowels but also consonants have formants, even though the formants of the latter differ from vowel formants and are less easily recognizable. This is due to the constriction at some point in the oral tract that is characteristic for the articulation of oral consonants. These constrictions cause antiresonances that eliminate or at least weaken the resonances at the same frequencies. The resulting shift of formant frequencies, the formant transitions, provide important perceptual cues with respect to manner of articulation (F1) and place of articulation (F2 and F3) of the respective consonant. Since neighboring consonants and vowels are known to influence each other, the exact shape of the formant transition is co-determined by the proximate vowel.

1.5 Structure of this book

The second chapter gives an overview of auditory and audiovisual speech perception in non-brain-damaged speakers and speakers with aphasia. In the third chapter, models of auditory and audiovisual speech perception are briefly sketched and evidence from neuroimaging studies for one or the other potential models is described. Furthermore, a study by Boersma (2006) is illustrated that tries to account for the McGurk effect in terms of Optimality Theory (OT). In the fourth chapter the first experiment of the current study is presented, describing the auditory word and non-word discrimination abilities of twelve Dutch aphasics. In the fifth chapter the McGurk effect is introduced and relevant information about the McGurk effect is provided. Based on this background the current
McGurk experiment was developed. It is described in detail in chapter six. The seventh chapter summarizes both experiments of the current study, discusses the main findings, and offers a prospect for further investigation.