The influence of phonetic features on aphasic speech perception

Dörte Hessler, Roel Jonkers & Roelien Bastiaanse

University of Groningen, Center for Language and Cognition Groningen

INTRODUCTION

Speech perception is multimodal. Not only auditory but also visual information (seen speech) is processed in perception (Rosenblum, 2008). It has been proven that seeing the speaker facilitates comprehension in a noisy environment or with demanding contents. Further evidence for the contribution of seen speech was added by the McGurk effect (McGurk and MacDonald, 1976). This effect describes the fusion between seen and heard speech. Subjects watched dubbed videos with auditory and visual information that did not match and were asked to report what they perceived. Instead of answering with the auditory (/ba/) or the visual (/ga/) component of the video they mostly reported a fusion of both (/da/). Furthermore, studies have shown that aphasic listeners benefit from seeing the speaker's face (Shindo, 1991). It is however unclear which phonetic features ("place of articulation", "manner of articulation" and/or "voicing") can actually make use of the additional information from seen speech.

Blumstein et al. (1977) showed that English-speaking aphasic listeners had more problems detecting differences in only one phonetic feature than differences in two or all three. Klitsch (2008) used the Dutch version of the PALPA (Bastiaanse et al., 1995) to investigate whether there were differences in detecting distinctions in the features "place of articulation", "manner of articulation" and "voicing". She came to the cautious conclusion that "place of articulation" was affected most, but also noted that the feature "voicing" could not be compared reliably to the other features, as the material included "voicing" contrasts only in initial, the other contrasts however only in final position or metathesis.

AIM

The current study investigates the influence of (additional) lip-reading on the aphasic subject's perception of speech. It furthermore aims at determining whether Dutch aphasic subjects also can detect wider distinctions (three phonetic features) more easily than more narrow ones. Finally we examine which phonetic features are most vulnerable in aphasia if manipulated in the same position.

PROCEDURE

So far 4 Dutch aphasic subjects with problems in comprehension as well as 14 non-brain-damaged control subjects participated in the current study. All aphasic subjects were at least 4 month post onset (4-146, average 43.5) and between 47 and 64 years (average 54) old. 1 was female.

A nonword discrimination task (same-different judgment) was carried out to investigate how different phonetic features were affected by aphasia. In this task there were three conditions: "auditory only" (subjects could only hear the speaker), "visual only" (the speaker could only be seen) and "audiovisual" (subjects heard and saw the speaker).

Material consisted of CVC(C) syllables that formed strings phonologically possible but non-existing in Dutch. The stimuli were recorded in a quiet environment and spoken by a male Dutch native speaker. All differences occurred in the first segment only because Dutch has final devoicing precluding a distinction in voicing in the final segment. The material was controlled for the amount of features differing within a pair as well as the actual feature differentiating it (see Figure 1).

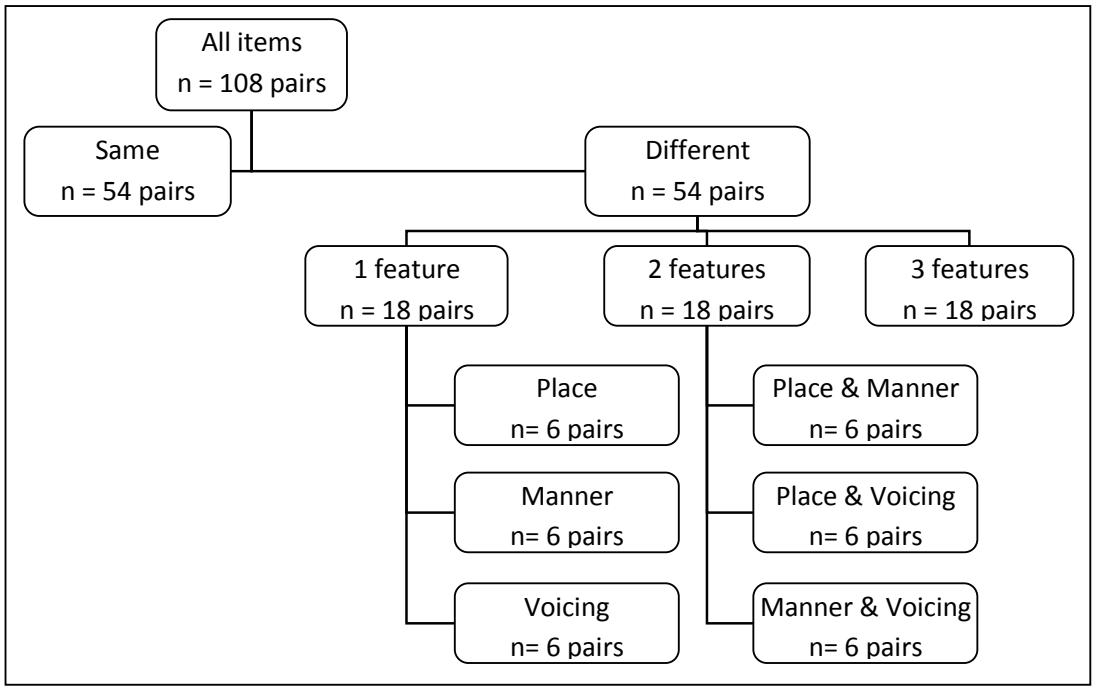

Figure 1: Outline of material used in discrimination task

RESULTS

Non-brain-damaged control subjects scored at ceiling for the "auditory only" and "audiovisual" conditions. In the "visual only" condition they performed worse, failing mainly in contrasts involving only "voicing" or "manner of articulation" or the combination of both.

Aphasic subjects scored significantly lower than control subjects on all conditions (see table 1). Their performance also differed between the three conditions (Friedman Anova: $X^2 = 8$, df = 2, $p < .05$). Performance was best in the "audiovisual" condition and worst in the "visual only" condition.

Condition	Controls (avg. correct)	Aphasic subj. (avg. correct)	Z-Score	p-value
Auditory only condition	99%	86%	-3.042	p = .002
Audiovisual condition	99%	90%	-3.120	p = .002
Visual only condition	83%	64%	-2.980	p = .003

Table 1: Comparison of aphasic and nbd-control group with Mann-Whitney U test (2-tailed)

Further analyses revealed that the number of features differing within the pair indeed plays a role for aphasic subjects in the "auditory only" (Friedman Anova: $X^2 = 6.857$, df = 2, $p < .05$) and "audiovisual" (Friedman Anova: $X^2 = 7.429$, df = 2, $p < .05$) conditions. In both conditions differences in only one feature are least often detected (67% for "auditory only" and 73% "audiovisual") compared to similar results for differences in two and three features ("auditory only": 89% and 87%; "audiovisual": 89% and 93%). Furthermore the influence of type of feature ("place" vs. "manner" vs. "voicing") was investigated, leading to significant results for the "audiovisual" condition (Friedman Anova: $X^2 = 6.533$, df = 2, $p < .05$) and marginally significant results for the "auditory only" condition (Friedman Anova: $X^2 = 5.692$, df = 2, $p = .058$). For both conditions it appears that "voicing" was the most difficult distinction, followed by "place of articulation" and "manner of articulation" (see Figure 2).

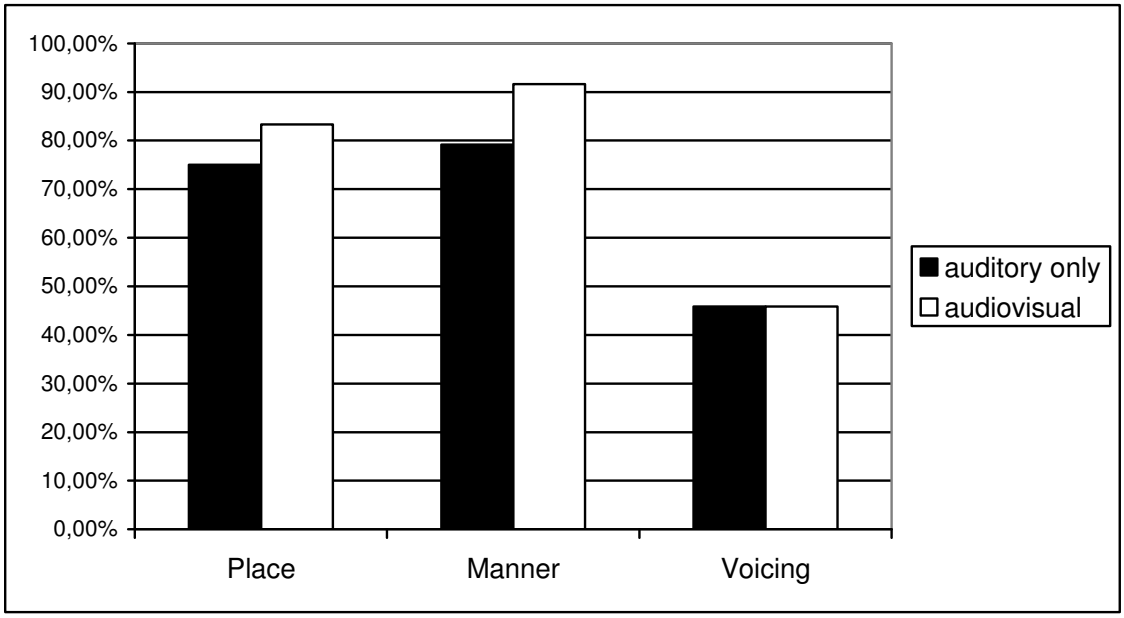

Figure 2: Percentage of correct aphasic responses to different features in auditory only and audiovisual conditions

DISCUSSION

The results obtained so far indicate that aphasic subjects have difficulties in discriminating pairs of nonwords. These difficulties are more profound for smaller differences. As previously reported by Blumstein et al. (1977) for English, we also find for Dutch speakers that differences are less likely to be detected if the items within the pair differ by only one phonetic feature. This holds for the "auditory only" as well as for the "audiovisual" condition.

Klitsch (2008) reported that "place of articulation" seems to be the most vulnerable feature; we found opposite results in the current study, showing that for the subjects included so far differences in "voicing" were least likely to be detected. This difference in results can be explained by the different materials used. In the PALPA (Bastiaanse et al., 1995) discrimination tasks, used by Klitsch (2008), the position of the difference is not balanced between "place of articulation", "manner of articulation" and "voicing". In the current study, however, all differences were manipulated in initial position providing a better basis for a comparison between the features.

The current study clearly shows that if contrasts between items result from more phonetic features, they are more easily detected. Furthermore, the type of feature differentiating items is of importance, indicating that differences in "voicing" are most difficult to perceive for Dutch aphasic listeners.

REFERENCES

Bastiaanse, R., Bosje, M., & Visch-Brink, E. (1995). *PALPA: Nederlandse Versie*. Hove: Lawrence Erlbaum.

Blumstein, S. E., Baker, E., & Goodglass, H. (1977). Phonological factors in auditory comprehension in aphasia. *Neuropsychologia, 15*(1), 19-30.

Klitsch, J. (2008). Open your eyes and listen carefully. Auditory and audiovisual speech perception and the McGurk effect in Dutch speakers with and without aphasia. *Groningen Dissertations in Linguistics (GRODIL), 67*.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748.

Rosenblum, L. D. (2008). Speech Perception as a Multimodal Phenomenon. *Current Directions in Psychological Science, 17*(6), 405-409.

Shindo, M., Kaga, K., & Tanaka, Y. (1991). Speech discrimination and lipreading in patients with word deafness or auditory agnosia. *Brain and Language 40*(2), 153-161.