Speed of Processing Phonological Information Presented Visually and Speechreading Proficiency

Jean-Pierre Gagné, Manuel Charbonneau, and Tony Leroux
Université de Montréal

An experiment was designed to investigate whether the speed of processing phonological information presented visually is related to speechreading proficiency even after the proportion of the variance associated with an analytical consonantal viseme recognition task was taken into consideration. Young adults ($n = 48$) with normal hearing and vision completed 4 different visual-speech perception tasks: (a) a sentence level speechreading test, (b) a visual-consonant recognition task with instructions to focus on optimizing performance level, (c) the same visual-consonant recognition task with instructions to respond as rapidly as possible to each test-item, and, (d) a visual-gating task. The mean response-time data and the results from the visual-gating task were significantly correlated to the performances on the sentence speechreading test, indicating that speed of processing at the phonological level is related to speechreading proficiency for sentences. Performance on the visual-gating task was a better predictor of speechreading performance than was response-time on the viseme-recognition task. Finally, hierarchical regression analyses revealed that together both the ability to identify consonants presented visually-only and the speed at which this task is accomplished contribute significantly to explaining the variance in performance on the speechreading test of sentence-length material.

Boothroyd (1988) has defined speechreading as: “the process of perceiving spoken language using vision as the sole source of sensory evidence” (p. 77). Some

---

Jean-Pierre Gagné, Manuel Charbonneau, and Tony Leroux, École d'orthophonie et d'audiologie, Université de Montréal.

Manuel Charbonneau is now at Audiology Department, Institut Raymond-Dewar, Montreal, Quebec, Canada.

Correspondence concerning this article should be addressed to Jean-Pierre Gagné, École d'orthophonie et d'audiologie, Université de Montréal, 7077 Park Avenue, Montreal, Quebec, Canada H3N 1X7. Telephone: 514-343-7458. Fax: 514-343-2115. E-mail: jean-pierre.gagne@umontreal.ca

of the characteristics of speech make speechreading a relatively difficult and effortful task. For example, visually some consonants are difficult (impossible) to distinguish from each other (e.g., homophenes such as /p/, /b/, and /m/). Some speech utterances are produced at the back of the oral cavity (e.g., /g/, /k/, /ŋ/) which makes it difficult to identify their precise place of articulation. Due to the effects of coarticulation, the shape of some speech sounds may change (or the speech elements may become difficult to visualize) when they are produced in different phonetic context (e.g., /uku/ vs. /iki/; /ara/ vs. /uru/). Furthermore, the speechreader is required to decode the visual-speech cues available at a very fast rate. Consequently, it is almost impossible to decode complex stimuli (such as sentences and continuing discourse) flawlessly based solely on the perception of visual-speech cues (e.g., Berger, 1972; Bernstein, 2006; Bernstein, Demorest, & Tucker, 2000; Jeffers & Barley, 1971; Lyxell, Rönnberg, Andersson, Andersson, & Samuelsson, 1998; Rönnberg, 1995, 2003; Rönnberg, Andersson, Andersson, et al., 1998; Rönnberg, Rudner, Foo, & Lunner, 2008; Rönnberg, Samuelsson, & Lyxell, 1998). Due to the complex nature of the task, speechreading may be conceived as the process of perceiving speech under difficult and effortful conditions (Rönnberg et al., 2008; Rudner & Rönnberg, 2008). Thus, there are substrates of visual speech perception that are shared by all forms of speech perception, regardless of the stimulus or the sensory modality in which the task is performed (Grant & Seitz, 1998; Rönnberg, 2003; Rönnberg et al., 2008; Rudner & Rönnberg, 2008).

The perceptual and cognitive processes associated with speechreading, and the intricate interaction that may exist among these processes, are not yet well understood. This report addresses one aspect of this issue. Specifically, we are interested in the relationship that may exist between the speed at which visually presented speech elements are encoded and speechreading proficiency for sentence-length material. Several investigators have reported that processing speed is an important component of speechreading. Lyxell and Holmberg (2000) stated that "The importance of a fast access to verbal information in speechreading is rather evident, as spoken stimuli are degraded and short-lived by nature" (p. 506).

Rönnberg and his colleagues reported that speed of processing speech information is associated with speechreading proficiency (Lyxell et al., 1996; Lyxell, Andersson, Borg, & Ohlsson, 2003; Lyxell & Holmberg, 2000; Lyxell & Rönnberg, 1991; Lyxell et al., 1998; Rönnberg, 1990, 1995). Significant correlations have been reported between performance on sentence-length speechreading material and speed of speech processing (e.g., Andersson, Lyxell, Rönnberg, & Spens, 2001; Lyxell, 1994; Lyxell & Andersson, 1998; Lyxell et al., 2003; Lyxell & Rönnberg, 1991, 1992; Lyxell et al.,1998; Rönnberg, 1990; Rönnberg et al., 1999). In his conceptual model of speechreading, Rönnberg (1995) listed processing speed along with word decoding ability and verbal inference making ability as the three main variables that have a direct impact on speechreading proficiency.

Measures of speed of processing have been obtained for different perceptual and cognitive processes associated with speechreading. Investigators have reported significant correlations between performances on sentence-length speechreading material and different tasks of processing speed including measures of phonological processing speed, various indices of lexical decision making speed, and speed of responses on rhyme judgment tasks (e.g., Andersson et al., 2001; Lyxell, 1994; Lyxell & Andersson, 1998; Lyxell et al., 2003; Lyxell & Rönnberg, 1991, 1992; Lyxell et al., 1998; Rönnberg, 1990; Rönnberg et al., 1999). There is a general consensus that, for most of the perceptual and cognitive processes involved, the speed at which speech is processed is a significant predictor of speech-perception proficiency (e.g., Andersson et al., 2001; Frauenfelder, Scholten, & Content, 2001; Marslen-Wilson, 1990, 1994; Pichora-Fuller & Singh, 2006; Wingfield, 1996; Wingfield & Tun, 2001). The present investigation addressed issues related to speed of processing visual-speech at the phonological level. Throughout this report phonological processing is operationally defined as the ability to assign a linguistic label to visual-speech elements produced by a talker.

The motivation for the study came from our attempt to explain the results of previous investigations related to speechreading proficiency. There exists a significant relationship between speechreading proficiency for sentence-length material and performance on analytical speechreading tasks such as the recognition of visual-consonants or consonantal visemes presented in nonsense syllables (Bernstein, 2006; Bernstein et al., 2000; Hanin, 1988; Rönnberg, 1995). Notwithstanding this finding, typically the results of investigations reveal that there is a wide range of performance on any given visual-sentence recognition task, even among the people who obtain similar scores on the viseme-recognition task. As a first approximation, it could be assumed that the people who obtain similar results on the viseme-recognition task have similar phonological encoding abilities. If this is so, then a pertinent question becomes: What can explain the substantial differences in performance on the visual-sentence recognition test that is observed among people with similar phonological encoding abilities?

Several explanations may account for this pattern of results. For example, significant correlations have been reported between speechreading proficiency and performance on high-level cognitive tasks that assess inference-making abilities (Bode, Nerbonne, & Sahlstrom, 1970; Lyxell & Rönnberg, 1987, 1989; Samuelsson & Rönnberg, 1991; Sanders & Coscarelli, 1970; Tatoul & Davidson, 1961; Williams, 1982). People with similar phonological processing skills may differ in their ability to make inferences based on the linguistic and contextual cues at their disposal (Lyxell & Rönnberg, 1989; Rönnberg, 1995). The present study was designed to investigate whether another variable, namely phonological encoding speed, may account for some of the variance in speechreading proficiency for sentence-length material. Our reasoning for choosing to address this issue was as follows. First, as mentioned above, measures of processing speed have been shown to be significant predictors of speechreading proficiency. Second, it is likely that in a phonological encoding task with no restriction on response time such as a consonantal viseme-recognition test, a negligible difference would exist between people who can perform the task rapidly and those who cannot. Furthermore, the speed at which phonological information is encoded may not be a critical factor when the speech stimuli to be processed are short, linguistically simple, and when the task is not too challenging (i.e., a close-set response paradigm using nonsense syllables). However, because of the limited capacity of working memory (Kahneman, 1973; Lavie, 1995) the speed at which visually presented phonological information is encoded may be important when the speech signal is long, linguistically complex, and presented at a rapid rate (e.g., sentences). Because of the continuous inflow of "new" information during sentence recognition tasks, the amount of linguistic information that may be held in working memory over a specific period of time is limited. Then, that sequence of speech is replaced by other utterances that must also be processed by the speechreader. The speed at which one is able to accomplish this task influences how much speech information is encoded. Within a fixed period of time, people who process phonological information efficiently will encode more speech elements than those who are not efficient. Hence, people proficient at encoding phonological information rapidly will have more speech information at their disposal as they access other perceptual and cognitive processes to decipher the spoken message (e.g., the information they can use to make inferences). Based on this reasoning, it is anticipated that there will be a significant relationship between the speed at which visually presented phonological information is encoded and speechreading proficiency for sentence-length material.

In addition to investigating the direct relationship between measures of phonological processing speed and speechreading proficiency, the present study was designed to also take into account the general phonological encoding abilities of the participants. The reasoning underlying this decision was as follows. Recall that a motivating factor underlying the investigation was to explain the variance in performance on speechreading sentences while taking into account the phonological encoding skills of the speechreaders. One way to address this issue is to first determine the amount of variance on a sentence speechreading test that is accounted for by performance on a phonological encoding task (e.g., a consonantal viseme recognition task). Then, as a second step, analyses can be conducted to ascertain whether performance on a phonological speed of processing task accounts for any additional variance observed on the sentence-level speechreading task. Specifically, in the present investigation the primary research question was: Does speed of processing at the phonological level explain any variance observed on sentence-level speechreading tasks, beyond the variance that can be explained by one's performance on a consonantal viseme-recognition task?

Several experimental paradigms can be employed to measure speed of processing speech. The approach most often used consists of measuring the response time of participants. It is assumed that in a time-monitored task people who are proficient at encoding speech information at the phonological level will be able to identify the syllables more rapidly and thus they will be able to respond more quickly than those who do not encode speech information efficiently. If we accept this premise, then a person's response time on a visual-consonant recognition task may be used as an indicator of processing speed at the phonological level (e.g., Larsby, Hällgren, Lyxell, & Arlinger, 2005; Lyxell et al., 1996; Rönnberg, 1990; Rudner & Rönnberg, 2008).

Another experimental approach that may be employed to measure speed of processing at the phonological level consists of using a gating task. The gating paradigm rests on the premise that at a specific moment in time, usually well before the stimulus is completely presented, a person gathers sufficient phonetic information to proceed to the recognition of the speech unit presented (Grosjean, 1980). Experimentally, gating consists of a time-gated speech recognition task. The paradigm investigates the time-course of speech recognition by presenting portions of a speech segment (e.g., usually syllables or words) repeatedly. Initially, only a short segment, known as a *gate,* is presented to the participant. Then, on subsequent trials, progressively longer speech segments (i.e., an accumulation of gates) that include the part of the signal to be recognized are presented. After each trial the participant must identify the stimulus presented. The least number of gates required to correctly recognize the speech element is defined as the Isolation Point (IP; Grosjean, 1996). It can be assumed that a person with the shortest IP processes speech more rapidly than someone who requires more gates before correctly identifying the speech element. If so, then a significant correlation should be observed between the IP scores and speechreading proficiency for sentences.

In the present investigation two different tasks were used to characterize speech processing speed at the phonological level. One task consisted of measuring the response-time in a visual-consonant recognition task. The other task consisted of a visual gating task. The results made it possible to examine the predictive power of each task as a measure of speed of processing visual-speech at the phonological level.

METHODS

Participants

A group of 48 young adults (16 males and 32 females) took part in the investigation. All of them used Québec-French as their primary language of everyday communication. The participants were between 18 and 40 years of age and the majority of them were undergraduate university-level students. Furthermore, all the participants had normal hearing acuity thresholds (i.e., ≤ 20 dB HL at test frequencies of 250, 500, 1000, 2000, and 4000 Hz, re: American National Standards Institute, 1996) and normal (or corrected normal) binocular visual acuity (i.e., 20/40 or better as measured with a Snellen chart; Strouse Watt, 2004). None of the participants had completed any formal training in speechreading. All the participants voluntarily signed a consent form prior to taking part in the investigation and they received a small monetary compensation for the time they devoted to the experiment.

Test Stimuli

Each participant completed four different visual-speech perception tasks (without any sound): (a) a sentence-level speechreading test, (b) a visual-consonant recognition test with specific instructions to focus on optimizing performance level, (c) a visual-consonant recognition test with specific instructions to respond as rapidly as possible to each test-item, and, (d) a visual-gating task.

Sentence level speechreading task. The test consisted of forms A and B of the Épreuve franco-québécoise de lecture labiale (Lalande, Lafleur, & Lacouture, 1989). Each of the 49 videotaped test items, spoken by one female talker, consisted of one sentence that was between 5 and 9 words in length. Each sentence is a syntactically and semantically appropriate sentence in French (e.g., “Les petits enfants sont à la plage”). The test was completed as a pencil and paper test. Specifically, after each test-item the participant was requested to write-down as much of the test sentence as possible. The participants were encouraged to guess even if they were not certain of the sentence that was presented. Before the test proper was administered, the participants completed eight practice trials (that did not appear in the test proper) to ensure that they understood the task.

Performance on this test was scored on the basis of the number of pre-selected keywords that were recognized correctly (Lalande et al., 1989). The percentage of keywords recognized correctly was calculated for each participant. Further, in order to produce a scale in which the size of the variance would be unrelated to the mean performance, a “rationalize” arcsine transform (RAU; Studebaker, 1985) was applied to the data. Performance on this test was considered to be a representative measure of general speechreading proficiency for sentences. Hence, the results obtained on this test constituted the criterion variable for the regression analyses that were conducted.

Visual-consonant recognition test with specific instructions to optimize accuracy. This test consisted of a computer administered version of the “Épreuve d'identification visuelle et catégorisation de consonnes en français québécois” (Jutras, Gagné, Picard, & Roy, 1998). The test includes 17 consonants used in Québec-French. Each consonant is presented in an intervocalic position to form a bi-syllabic nonsense word (e.g., /a-C̲-a/). In the computer version of this test, the testing procedure is as follows: (a) a female-talker appears on the computer monitor (a 17 in. Viewsonic monitor, model P75f+) and produces one test-item, (b) a response-screen listing the 17 different consonants (written orthographically) appears on the computer monitor, (c) the participant uses the computer-mouse to indicate which consonant was presented and to record the response time, and (d) the participant presses a validation-button which is used to trigger the presentation of the following test-item. In this test, the same 17 bi-syllabic words are presented five times in a random order, for a total of 85 test-items. The participants were told that a response had to be provided for each test-item and they were instructed to guess if they were not certain of the correct response. Most importantly, in this version of the test administration, the participants were instructed to take as much time as required and to focus on providing the most accurate response possible. Before the test proper was administered, the participants completed eight-practice trials to ensure that they understood the task. For each participant, the percentage of correct visual-consonant recognition responses and the percentage of consonantal viseme recognition responses were calculated. For the latter score, the consonantal groupings (the visemes) used were those established for the video recording of the talker (see Jutras et al., 1998). Both the percentage visual-consonant recognition scores and the percentage viseme-recognition scores were converted into RAUs.

Visual-consonant recognition test with specific instructions to respond as rapidly as possible. This test consisted of the same test as the one described above. Except for the instructions given to the participants, the test administration procedures were identical as those described above. In this task, although the participants were encouraged to identify the test-item correctly they were instructed to respond as rapidly as possible after the presentation of the test-item. Once the response button was pressed the computer program stored the response provided as well as the time required to provide a response. Before the test proper was administered, the participants completed eight-practice trials to ensure that they understood the task. For each participant, the mean response time was computed. The mean response time was based on the response times obtained for the test-items for which the participant obtained a correct viseme-recognition score.

The visual-gating task. Six test stimuli taken from the Visual Consonant Recognition Test (Jutras et al., 1998) were selected for the visual-gating task. Each test stimulus consisted of a voiced consonant with a unique place of articulation (i.e., /b/, /d/, /g/, /ʒ/, /v/, and /z/). The following editing procedures were applied to the digital files in order to generate the gated stimuli. For each of the six bisyllabic nonsense words, the standard stimulus consisted of the video clip of the talker uttering the first part of the bisyllable. Specifically, that clip showed the talker from the beginning of the utterance to the video frame that occurred 165 ms before the point at which the talker reached a maximum mouth opening position while articulating the initial vowel (/a/) of the test stimulus. The visual gates were added from the end point of the standard stimulus. Specifically, 20 gates were added to each of the six test stimuli. Each gate was 33 ms in duration. Thus, for each of the six consonants, the first gate consisted of the standard stimulus plus the first 33 ms of the remainder of the test stimulus; the second gate consisted of the standard stimulus plus 66 ms of the remainder of the test stimulus, and so forth. . . . The final segment consisted of the standard stimulus plus 660 ms of the remainder of the test stimulus. In total, 480 gated test-items were shown to each participant (6 test stimuli × 20 gates × 4 repetitions of each gate).

The visual-gating task was computer administered. The test procedure was as follows: (a) a gated test-item was presented (visually only), (b) the participant was requested to use the computer-mouse to indicate on the response-screen which one of the six test stimuli had been presented, and (c) the participant clicked on the validation key to indicate that a final response had been provided. Then, the next gated test-item was presented. The test-items were presented in a randomized order (across consonants and gate durations). Before the test proper was administered, the participants completed eight-practice trials to ensure that they understood the task.

For each participant, and each test-stimulus, an IP (in number of gates) was determined. Two criteria were used to determine the IP: (a) the shortest gate at which the participant correctly identified the test-stimulus on at least 75% of the trials, and (b) the response rate did not drop below 75% for any of the longer gates. An IP of 21 gates was arbitrarily attributed to a participant who had not reached the established criteria for a given test stimulus even once the 20 gated stimuli had been presented.

During the data analysis, it was observed that an IP of 21 gates was attributed on 36 occasions (from a total of 288 possibilities), when all six test stimuli were considered. A substantial proportion of these occasions (i.e., 10/36) was associated with the recognition of the stimulus /aga/. Specifically, the results of the visual-gating task revealed that only 46% of the participants satisfied the response-criteria for the test-stimulus /aga/. The results obtained for this test-stimulus differed considerably from responses obtained for the other test-stimuli where it was possible to calculate a precise IP score for almost all of the participants (e.g., 98% of the participants for /aba/; 94% for /ada/; 98% for /aʒa/; 98% for /ava/; 92% for /aza/). Based on these results it was decided that the data obtained for the test-stimulus /aga/ would be discarded from the data analyses.

The responses obtained from the visual-gating task were used to calculate a “normalized IP score” for each participant. First, for each of the five test-stimuli, a z-score was computed for every participant. Then, for each participant, the z-scores obtained for each of the 5 test-stimuli were used to calculate a mean normalized IP score. This latter value was used to characterize a participant’s performance on the gating task.

Table 1
Means, Standard Deviations, and Range of Performances
on the Tests Administered

	M	*SD*	Range
Sentence speechreading task (RAU)	58.57	23.880	- 5.29 - 100.40
Visual-consonant recognition scores (RAU)	42.55	6.490	30.52 - 55.96
Consonantal viseme scores (RAU)	97.07	11.150	70.67 - 117.98
Visual-consonant recognition response-time (seconds)	2.05	.873	1.02 - 6.33
Mean normalized IP from visual gating task	0.00	2.100	-2.02 - 7.18

Note. RAU = rationalized arcsine unit. IP = isolation point.

General Testing Procedures

The complete test protocol was generally administered over two sessions of approximately 1 hr each. The test sessions were conducted in a quiet and visually non-distracting laboratory. The participants were tested individually. Each participant was seated at a distance of 0.5 m directly in front of the computer monitor and the computer mouse was set on a small table on the right-hand side of the participant. A Latin-square design was used to determine in which order the participants completed the four experimental tasks.

RESULTS

Descriptive statistics summarizing the results obtained on each visual speech-perception task are shown in Table 1. The relationships observed amongst each of those variables are displayed in Table 2. Not surprisingly, significant positive correlations were obtained between the performance on the visual-consonant recognition task that measured the accuracy of responses (both the viseme and the visual-consonant recognition scores) and performance on the sentence speechreading test. Also, there was a significant correlation between the performances on the visual-consonant recognition task that measured response-time and the sentence speechreading test as well as between the visual-gating task and the sentence test. It is noteworthy that there was a significant, but not perfect correlation between the normalized IP scores and the responses on the visual-consonant recognition task in which the response-time was measured. The results indicate that the scores from the visual-gating task constitute a better predictor of speechreading proficiency for sentences than the response-time data.

Table 2

Correlations Among the Variables Used in the Present Investigation

	Sentence speechreading scores	**Visual-consonant recognition scores**	**Consonantal viseme scores**	**Visual-consonant recognition response-time**	**Normalized IP scores from visual gating tasks**
Sentence speechreading scores	–	.401 ($p = .005$)	.5850 ($p < .0001$)	-.325 ($p = .024$)	-.6370 ($p < .0001$)
Visual-consonant recognition scores		–	.5940 ($p < .0001$)	-.020 ($p = .895$)	-.3290 ($p = .0230$)
Consonantal viseme scores			–	-.167 ($p = .258$)	-.4320 ($p = .0020$)
Visual-consonant recognition response time				–	.5910 ($p < .0001$)
Normalized IP scores from visual gating tasks					–

Note. IP = isolation point.

Hierarchical regression analyses were conducted to investigate whether each of the speed of processing measures would explain any of the variance on the sentence speechreading recognition test beyond the variance accounted for by the performance on an analytical visual-consonant recognition tasks. Two separate series of hierarchical regression analyses were performed with the normalized-IP data. In one analysis, the first predictor variable entered into the regression model consisted of the viseme score and the second predictor variable consisted of the Mean IP score from the visual-gating task (see Table 3, line A). The results revealed that each of the two predictor variables contributed significantly to explaining a total of 52.4% of the variance on the sentence speechreading test. A different hierarchical regression analysis was computed using the visual-consonant recognition scores and the mean normalized-IP scores as the predictor variables (see Table 3, line B). The results of this analysis revealed that combined those two variables contributed to explaining 44.7% of the variance observed on the sentence speechreading test.

Similarly, two separate hierarchical analyses were computed with the accuracy results on the visual-consonant recognition task and the response-time data on the same task. In one analysis, the first predictor variable entered into the regression analysis consisted of the accuracy scores on the viseme task and the second predictor variable consisted of the response-time on the visual-consonantal recognition task (see Table 3, line C). The additional variance explained by including the response-time data was not statistically significant. A fourth hierarchical regression analysis was computed using the accuracy scores as the first predictor variable and the response-time data as the second predictor variable (see Table 3, line D). The results revealed that together these two variables contributed significantly to explain 27.2% of the variance observed on the sentence speechreading test.

DISCUSSION

The results are consistent with those of previous studies that used a response-time measure to characterize speed of processing. The response-time data explained 11% of the variance observed on the sentence speechreading test. Rönnberg (1995) reported that the general speed with which visual decoding, lexical access, and processing of linguistic input is carried out generally accounts for approximately 15% of the variance observed on sentence-level speechreading tasks. The normalized-IP scores explained 41% of the variance observed on the sentence speechreading test. Further investigations are required to quantify the true unexplained variance associated with the normalized-IP scores as well as the response-time data. Because reliability data were not collected for either of those predictor variables it is not possible to assess the error variance associated with each task.

The normalized-IP scores explained substantially more of the variance ob-

Table 3

Results of Hierarchical Regression Analyses

First variable entered hierarchical regression analysis	R^2 for 1st variable	Significance for 1st variable	2nd variable entered hierarchical regression analysis	R^2_{change}	Significance of R^2_{change}	R^2_{total}
A. Consonantal viseme scores	.342	< .0001	Mean normalize IP from visual gating task	.182	< .0001	.524
B. Visual consonant recognition scores	.161	< .0050	Mean normalize IP from visual gating task	.286	< .0001	.447
C. Consonantal viseme scores	.342	< .0001	Visual-consonant recognition response-time	.053	.0520	–[a]
D. Visual consonant recognition scores	.161	< .0050	Visual-consonant recognition response-time	.111	.0120	.272

Note. IP = isolation point.

[a] The additional variability explained by including the Mean Response-Time data is not statistically significant.

served on the sentence speechreading test than the response-time data. One possible explanation for this result may be that the inter-participant variability (i.e., the standard deviation) and the range of scores were greater for the visual-gating task than for the response-time data. Regression analyses involve the computation of correlation coefficients which are sensitive to the dispersion of scores obtained by the group of participants. The larger individual differences observed on the visual-gating task may explain why this variable constituted a more discriminative predictor of speech processing speed proficiency.

It is possible that, although the same bisyllables were used in both the visual-gating task and the visual-consonant recognition tasks, more importance was attributed to the speech recognition component of the task in the gating task than in the visual-consonant recognition response-time task. In the visual-gating task the focus is clearly on identifying the visual-speech element presented. However, in the response-time task the speech recognition component of the procedure may be considered as being secondary given that the instructions emphasize the fact that the participant is expected to provide a response as rapidly as possible. Hence, the participant may devote more overall resources to identifying the speech element correctly in the gating task than in the response-time task. From this perspective the results obtained from the visual-gating task may constitute a better indicator of speed of processing speech at the phonological level. Furthermore, the response-time task was a more complex task than the gating task. In the visual-consonant recognition task used to measure response-time, the response had to be selected among a set of 17 bisyllabic stimuli. In the visual-gating task the correct response had to be selected from a set of only six possible choices.

Although both procedures provide an index of speed of processing there are important differences between the two experimental tasks used in the present investigation. On the one hand, a response-time measurement holds a high level of validity because it constitutes a direct measure of processing speed. That is, it consists of a measure of the time taken by a participant to respond after a stimulus has been presented. However, due to the nature of the task the responses obtained from this type of task are contaminated by time-related factors other than speech processing speed per se. For example, individual differences in the time required to identify the designated area to indicate the response on the computer screen, the time required to move the computer-mouse to the intended position on the screen, and other time related factors contribute to the error of measurement obtained when a response-time measure is used as an independent variable. Related to construct validity, results of gating experiments are less appealing than response-time paradigms because the participant's response provides an indirect measure of speed of processing. That is, it is assumed that people who require fewer gates to recognize a speech element, process speech more rapidly than those who require more gates before they are able to identify the speech element.

However, one cannot exclude that some higher level cognitive processes (e.g., the ability to make inferences based on one's knowledge of visual articulatory trajectories of consonants) can be accessed during a gating task. Specifically, the data from the gated task reflect the participant's ability to identify segments of consonants on the basis of incomplete information (i.e., missing gates). This ability may, by itself influence the scores obtained on the gating task. It is noteworthy that performance on the sentence task is also a measure of the observer's ability to identify speech segments on the basis of incomplete information (due to intrinsic nature of speechreading sentence-length material). Thus, the ability to infer a signal from incomplete information, and the speed at which this task is accomplished, is applied both in the gating task and in the visual-sentence recognition task. This may explain, at least partly, the strength of the statistically significant correlation between the visual-gating task and performance on the sentence speechreading task. Notwithstanding the involvement of higher cognitive processes one significant advantage of using a gating paradigm is that the data are not contaminated by errors of measurement attributable to differences in the time required by the participants to actually perform the experimental task. Although the procedure measures aspects of speed of processing speech, the actual time required to respond does not enter into the response provided. The present results do not make it possible to specify the different aspects of speed of processing that are measured with the gating task and in the visual-consonant recognition task that was used to measure response-time. Future investigations should address this issue.

One goal of the investigation was to determine whether speed of processing visually presented speech information at the phonological level would explain a significant portion of the variance observed on a sentence level speechreading test even after the variance associated with one's performance on an analytic visual-consonant recognition task was taken into account. The hierarchical regression analyses revealed that, once the proportion of the variance associated with the performances on the viseme-recognition task was accounted for, performances on the speed of processing tasks contributed significantly to explaining additional variance observed on the sentence-length speechreading test. The results were more convincing for the normalized IP data than for the response-time data. Empirically, this finding can be explained in part by the stronger correlation obtained between the results of the visual-gating task and speechreading proficiency for sentence-length material than the correlation observed between the response-time data and speechreading proficiency for sentences. Conceptually, the results indicate that both the general ability to encode speech elements at the phonological level as well as the speed with which this is accomplished contribute independently to explaining the variance in speechreading proficiency for sentences. This finding supports Rönnberg's model of speechreading whereby he identified word decoding ability and speed of processing as two variables that have a direct impact on speechreading proficiency for sentences (Rönnberg, 1995).

When they were considered together, at best performance on a general phonological encoding task and performance obtained on a speed of processing task at the phonological level explained approximately 52% of all of the variance observed on the speechreading task involving sentence-length material. Investigations are needed to identify other perceptual and cognitive processes that may account for the remaining unexplained variance observed on sentence speechreading tests. Considering the speechreading literature as well as our current understanding of speech perception under effortful conditions (in any perceptual modality), at least two candidates should be considered. First, several studies have shown that the ability to make inferences and to apply contextual and linguistic information to speechreading is significantly correlated with speechreading proficiency (Bode et al., 1970; Lyxell & Rönnberg, 1987, 1989; Samuelsson & Rönnberg, 1991; Sanders & Coscarelli, 1970). Second, several investigators have observed a significant relationship between working memory span and speechreading scores for sentence-length material (e.g., Lyxell & Rönnberg, 1992; Rönnberg, 1995; Rönnberg et al., 2008; Rudner, Foo, Lunner, & Rönnberg, 2008). It would be of interest to ascertain whether these higher-level cognitive skills would explain any of the variance on a test of speechreading proficiency, once the variance associated with analytical visual-consonant recognition ability and speed of processing at the phonological level skills were taken into account.

ACKNOWLEDGEMENTS

This research grant was supported in part by a research grant awarded to the first author (Jean-Pierre Gagné) by the National Science and Engineering Research Council (NSERC). The authors would like to thank Jean-François Angers and Miguel Chagnon from de Département de Mathématiques et Statistiques at the Université de Montréal for their contribution to the data analysis.

REFERENCES

American National Standards Institute. (1996). *American national standard specification for audiometers* (ANSI S3.6-1996). New York: Author.

Andersson, U., Lyxell, J., Rönnberg, J., & Spens, K.E. (2001). Cognitive correlates of visual speech understanding in hearing-impaired individuals. *Journal of Deaf Studies and Deaf Education, 6,* 103-115.

Berger, K.W. (1972). *Speechreading: Principles and methods.* Baltimore: National Educational Press.

Bernstein, L.E. (2006). *Visual speech perception.* Unpublished manuscript.

Bernstein, L.E., Demorest, M.E., & Tucker, P.E. (2000). Speech perception without hearing. *Perception and Psychophysics, 62,* 233-252.

Bode, D.L., Nerbonne, G.P., & Sahlstrom, L.J. (1970). Speechreading and the synthesis of distorted printed sentences. *Journal of Speech and Hearing Research, 19,* 115-121.

Boothroyd, A. (1988). Linguistic factors of speechreading. In C.L. De Filippo & D.G. Sims (Eds.), *New Reflections on Speechreading* [Monograph]. *The Volta Review, 90,* 77-87.

Frauenfelder, U., Scholten, M., & Content, A. (2001). Bottom-up inhibition in lexical selection: Phonological mismatch effects in spoken word recognition. *Language and Cognitive Processes, 16,* 583-607.

Grant, K.W., & Seitz, P.F. (1998). Measures of auditory-visual integration in nonsense syllables and sentences. *Journal of the Acoustical Society of America, 104,* 2438-2450.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics, 28,* 267-283.

Grosjean, F. (1996). Gating. *Language and Cognitive Processes, 11,* 597-604.

Hanin, L. (1988). *The effect of experience and linguistic context on speechreading.* Unpublished doctoral dissertation, The City University of New York.

Jeffers, J., & Barley, M. (1971). *Speechreading (lipreading).* Springfield, IL: Thomas Publishers.

Jutras, B., Gagné, J.-P., Picard, M., & Roy, J. (1998). Identification visuelle et catégorisation de consonnes en français québécois (Visual identification and categorization of Québec-French consonants). *Journal of Speech-Language Pathology and Audiology, 22,* 81-87.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Lalande, N.M., Lafleur, G., & Lacouture, Y.S. (1989). Développement d'une épreuve franco-québecoise de lecture labiale (Development of a speechreading test in Québec-French). *Audiology, 28,* 71-81.

Larsby, B., Hällgren, M., Lyxell, B., & Arlinger, S. (2005). Cognitive performance and perceived effort in speech processing tasks: Effects of different noise backgrounds in normal-hearing and hearing-impaired subjects. *International Journal of Audiology, 44,* 131-143.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology, 21*(3), 451-468.

Lyxell, B. (1994). Skilled speechreading: A single-case study. *Scandinavian Journal of Psychology, 35,* 212-219.

Lyxell, B., & Andersson, U. (1998). Phonological capabilities and speech understanding. *Proceedings of AVSP: International conference on auditory-visual speech processing, Australia, 1998,* 71-174.

Lyxell, B., Andersson, J., Arlinger, S., Bredberg, G., Harder, H., & Rönnberg, J. (1996). Verbal information-processing capabilities and cochlear implants: Implications for preoperative predictors of speech understanding. *Journal of Deaf Studies and Deaf Education, 1,* 190-201.

Lyxell, B., Andersson, U., Borg, E., & Ohlsson, I.S. (2003). Working memory capacity and phonological processing in deafened adults and individuals with a severe hearing impairment. *International Audiology, 42*(Suppl. 1), S86-S89.

Lyxell, B., & Holmberg, I. (2000). Visual speechreading and cognitive performance in hearing-impaired children and normal hearing children. *British Journal of Educational Psychology, 70,* 505-518.

Lyxell, B., & Rönnberg, J. (1987). Guessing and speechreading. *British Journal of Audiology, 21,* 13-20.

Lyxell, B., & Rönnberg, J. (1989). Information-processing skills and speechreading. *British Journal of Audiology, 23,* 339-347.

Lyxell, B., & Rönnberg, J. (1991). Visual speech processing: Relations to sentence-based speechreading and hearing-impairments. *Scandinavian Journal of Psychology, 32,* 9-17.

Lyxell, B., & Rönnberg, J. (1992). Verbal ability and sentence-based speechreading. *Scandinavian Audiology, 21,* 67-72.

Lyxell, B., Rönnberg, J., Andersson, U., Andersson, J., & Samuelsson, S. (1998). Working memory and visual speech communication. *Proceedings of AVSP: International conference on auditory-visual speech processing, Australia, 1998,* 127-130.

Marslen-Wilson, W.D. (1990). Activation, competition, and frequency in lexical access. In G.T.M. Altman (Ed.), *Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 148-172). Cambridge, MA: MIT Press.

Marslen-Wilson,W.D. (1994). Speech perception as a cognitive process: The role of abstractness and inference. *Proceeding of the 1994 International Conference on Spoken Language Processing*. Yokohama, Japan: Acoustical Society of Japan.

Pichora-Fuller, M.K., & Singh, G. (2006). Effects of age on auditory and cognitive processing: Implications for hearing aid fitting and audiologic rehabilitation. *Trends in Amplification, 10,* 29-59.

Rönnberg, J. (1990). Cognitive and communicative functions: The effects of chronological ages and "handicap age." *European Journal of Cognitive Psychology, 2,* 253-273.

Rönnberg, J. (1995). What makes a skilled speechreader? In G. Plant & K.E. Spens (Eds.), *Profound deafness and speech communication* (pp. 393-416). London: Whurr Publishers Ltd.

Rönnberg, J. (2003). Cognition in the hearing impaired and deaf as a bridge between signal and dialogue: A framework and a model. *International Audiology, 42* (Suppl. 1), S68-S76.

Rönnberg, J., Andersson, J., Andersson, U., Johansson, K., Lyxell, B., & Samuelsson, S. (1998). Cognition as a bridge between signal and dialogue: Communication in the hearing-impaired and deaf. *Scandinavian Audiology, 27* (Suppl. 49), 101-108.

Rönnberg, J., Andersson, J., Samuelsson, S., Söderfeldt, B., Lyxell, B., & Risberg, J. (1999). A speechreading expert: The case of MM. *Journal of Speech, Language and Hearing Research, 42,* 5-20.

Rönnberg, J., Rudner, M., Foo, C., & Lunner, T. (2008). Cognition counts: A working memory system for ease of language understanding (ELU). *International Journal of Audiology, 47* (Suppl. 2), S99-S105.

Rönnberg, J., Samuelsson, S., & Lyxell, B. (1998). Conceptual constraints in sentence-based lipreading in the hearing-impaired. In R. Campbell, B. Dodd, & D. Burnham (Eds.), *Hearing by eye II: Advances in the psychology of speechreading and auditory-visual speech* (pp. 143-153). Hove, UK: Psychology Press Ltd.

Rudner, M., Foo, C., Lunner, T., & Rönnberg, J. (2008). Phonological mismatch and explicit cognitive processing in a sample of 102 hearing aid users. *International Journal of Audiology, 47* (Suppl. 2), S91-S98.

Rudner, M., & Rönnberg, J. (2008). The role of the episodic buffer in working memory for language processing. *Cognitive Processing, 9,* 19-28.

Samuelsson, S., & Rönnberg, J. (1991). Script activation in lipreading. *Scandinavian Journal of Psychology, 32,* 124-143.

Sanders, J.W., & Coscarelli, J.E. (1970). Relationship of visual synthesis skill to lipreading performance. *American Annals of the Deaf, 115,* 23-25.

Strouse Watt, W. (2004). *How visual acuity is measured?* Accessed April 17, 2007, from http://www.mdsupport.org/library/acuity.html

Studebaker, G.A. (1985). A «rationalized» arcsine transform. *Journal of Speech and Hearing Research, 28,* 455-462.

Tatoul, C.M., & Davidson, G.D. (1961). Lipreading and letter prediction. *Journal of Speech and Hearing Research, 4,* 178-181.

Williams, A. (1982). The relationship between two visual communication systems: Reading and lipreading. *Journal of Speech and Hearing Research, 25,* 500-503.

Wingfield, A. (1996). Cognitive factors in auditory performance: Context, speed of processing, and constraints of memory. *Journal of the American Academy of Audiology, 7,* 175-182.

Wingfield, A., & Tun, P. (2001). Spoken language comprehension in older adults: Interaction between sensory and cognitive change in normal aging. *Seminars in Hearing, 22,* 287-301.