Shared lexical access processes in speaking and listening? An individual differences study

Florian Hintz,[1*] Suzanne R. Jongman,[1*] Marjolijn Dijkhuis[1], Vera van t' Hoff[1], James M. McQueen[1,2],
Antje S. Meyer, [1,2]
[1]*Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands*
[2]*Radboud University, Nijmegen, The Netherlands*
*Both authors contributed equally

------In press at Journal of Experimental Psychology: Learning, Memory, and Cognition------

Running title: Lexical access in picture naming and lexical decision

Keywords: individual differences; lexical decision; picture naming; processing speed

Correspondence should be addressed to:
Florian Hintz
Max Planck Institute for Psycholinguistics
P.O. Box 310
6500 AH Nijmegen
The Netherlands
Telephone: +31-24-3521335
Fax: +31-24-3521213
florian.hintz@mpi.nl

Abstract

Lexical access is a core component of word processing. In order to produce or comprehend a word, language users must access word forms in their mental lexicon. However, despite its involvement in both tasks, previous research has often studied lexical access in either production or comprehension alone. Therefore, it is unknown to which extent lexical access processes are shared across both tasks. Picture naming and auditory lexical decision are considered good tools for studying lexical access. Both of them are speeded tasks. Given these commonalities, another open question concerns the involvement of general cognitive abilities (e.g., processing speed) in both linguistic tasks. In the present study, we addressed these questions. We tested a large group of young adults enrolled in academic and vocational courses. Participants completed picture naming and auditory lexical decision tasks as well as a battery of tests assessing non-verbal processing speed, vocabulary, and non-verbal intelligence. Our results suggest that the lexical access processes involved in picture naming and lexical decision are related but less closely than one might have thought. Moreover, reaction times in picture naming and lexical decision depended as least as much on general processing speed as on domain-specific linguistic processes (i.e., lexical access processes).

Key words: processing speed, picture naming, lexical decision, individual differences

Shared lexical access processes in speaking and listening? An individual differences study

Introduction

Experimental psycholinguistics has primarily been concerned with general processing principles expected to apply to all adult speakers of a language. This research has been highly successful, as it has led to the formulation and empirical evaluation of detailed models of speaking and listening, (Hagoort, in press). Recently, this research tradition has been complemented by work addressing individual differences in language skills among native speakers (e.g., Dabrowska, 2018; Welcome & Joanisse, 2014; see also Kidd, Donnelly, & Christiansen, 2017). As several authors have pointed out, comprehensive psychological models should not only explain average or modal behavior, but also the spread around the central tendency (Andrews, 2012; Engelhardt et al., 2017; Welcome & Joanisse, 2014). Relatedly, models capturing the behavior of university students – the typical participants in most psycholinguistic studies – may or may not apply to broader samples, including, for instance, persons with different educational backgrounds (e.g., Adelman, Sabatos-deVito, Marquis, & Estes, 2014).

Other reasons for the growing interest in individual differences are more specific to psycholinguistics. One of them is the strong consensus in the field that utterance production and comprehension are lexically driven incremental processes, that is, they hinge on the swift access to information stored in the mental lexicon (e.g., Bock & Ferreira, 2014; Chater, McCauley, & Christiansen, 2016; Konopka & Meyer, 2014; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). As adults are likely to differ in lexical knowledge (the number and types of words they know), considerations of individual differences in language skills naturally move into focus. Moreover, there is accumulating strong evidence that the language system is not a module that can be separated from other components of the cognitive system (Anderson, Chiu, Huette, & Spivey, 2011; Engelhardt et al., 2017). Instead, using language always involves other cognitive components, including, for instance, visual, motor, attentional, and memory processes (e.g., McQueen & Meyer, in press). As adults are highly likely to differ in domain-general cognitive skills, they are bound to differ in the way they use language as well. Therefore, comprehensive theories of speaking and listening must capture how different components of the linguistic system work together when language is produced or understood, and how individual differences in relevant non-linguistic abilities or skills affect utterance generation and comprehension. Such comprehensive theories are essential for giving sound practical recommendations for language testing and instruction.

Studies that examine the degree to which task performance varies across individuals can contribute to the development of these comprehensive theories. In particular, such studies can indicate which cognitive components (specific processes or representations) are shared across tasks. For example, analysis revealing substantial amounts of shared variance between measures of speaking and listening performance would point to a shared underlying cognitive component. To develop a comprehensive account of language processing, it would be necessary to carry out such analyses for all aspects of language use. The current paper is less ambitious. We focus on lexical access, a process that is essential to both speech production and speech comprehension. Broadly defined, lexical access is the process of accessing representations of words in the mental lexicon. We took an individual differences approach to ask how much commonality there is between the lexical access processes involved in speech production and those involved in speech comprehension. Since we used response latency measures (in picture naming and auditory lexical decision, respectively), it was important to also determine the participants' general processing speed. This allowed us to ask how strongly the correlation between the latencies in the linguistic tasks was moderated by processing speed and hence to address the degree to which commonalities across the linguistic tasks are the result of repeated use of a domain-general cognitive skill rather than of the process of lexical access. Furthermore, we also obtained indicators of their vocabulary size and general intelligence to be able to factor out variance explained by these two variables.

In sum, the goals of this project were to determine (1) how strongly the latencies of the two linguistic tasks correlated with each other, thereby assessing the extent to which lexical access processes are shared between production and comprehension, and (2) how strongly the correlation between the latencies in the linguistic tasks was moderated by processing speed, that is, the extent to which linguistic processing involved domain-general abilities.

The choice of linguistic and non-linguistic tasks

We used picture naming because it has been widely used before in language production research as a tool for studying lexical access during speaking. Based on a large body of work using this task, there is good consensus about the main processing steps occurring in picture naming and their temporal coordination. Briefly, speakers need to identify the concept represented in the picture, activate its lexical concept, select its lemma, encode its morphological and phonological form and, finally, encode its phonetic form (Levelt, Roelofs, & Meyer, 1999). These processes may overlap in time (e.g. Indefrey & Levelt, 2004; Strijkers et al., 2009). The auditory lexical decision task plays a similar key role in studies of lexical access in spoken word comprehension. To decide whether or not a sound sequence is a word of their language, listeners must identify the incoming phonetic representations, attempt to map it onto a phonological representation stored in their mental lexicon and, depending on whether or not a sufficient match can be found within a set time period, indicate the decision through a button press (Goldinger, 1996; Ratcliff & McKoon, 1997). In the literature, the visual version of the lexical decision task has probably been used more frequently than the auditory version. We opted for the latter version because our goal was to understand spoken language processing and because we aimed to avoid any direct influences of participants' technical reading skills on their performance. In order to facilitate the interpretation of the results, the names of the objects used in the picture naming task and the target words in the lexical decision task were matched for frequency and neighborhood density, variables that have been shown to affect the latencies in many studies (Jescheniak & Levelt, 1994; Luce & Pisoni, 1998; Peramunage et al., 2011). Another control was prevalence (i.e., the degree to which a given word is known by a representative sample of the Dutch speaking population, Keuleers et al., 2015). We chose common words that were known by at least 97% of the population to increase name agreement in the picture naming and to increase the likelihood of a Yes-response to word stimuli in the auditory lexical decision task.

As stated above, our main research question was to assess how strongly performance indicators (i.e., response latencies) in the two tasks would be correlated. While each of the two tasks has been used in numerous studies, we know of only one study that has used them together (Litcofsky, Tanner, & van Hell, 2015). In that study, bilingual participants (N=42) were tested in their first and second language, and moderate correlations between production and comprehension scores were found for both languages. However, only accuracy but not response speed was measured. In the present study, we measured accuracy and response speed in the participants' first language. High cross-task correlations would indicate that the two tasks recruit shared cognitive processes, whereas low correlations would indicate that the processes involved in each task have little in common.

Which cognitive processes might be involved in picture naming and lexical decision? In both tasks, participants must access word form information in their mental lexicon[1]. To the extent that the processes involved in lexical access for speaking (picture naming in particular) and comprehension (auditory lexical decision in particular) are related, one should find a correlation. However, lexical decision and picture naming also have in common that they are both timed tasks. Therefore, a correlation in lexical decision and picture naming speed could arise because both tasks tap domain-general cognitive speed. One might predict that persons who are fast in carrying out cognitive tasks in general (or at least in a lab environment) should also be fast in picture naming and lexical decision. Alternatively, one could argue that performing linguistic tasks involves rather specialized skills and that no strong relationship to general cognitive speed is to be expected. For example, if participants carried out the picture naming and lexical decision tasks in a second language, their proficiency in that language might predict their performance much better than their processing speed. Similarly, when the tasks are carried out in the participants' first language, their performance might likewise hinge more on specific linguistic skills than on their general processing speed.

The second goal of the present study was to assess how strongly processing speed moderated the correlation between the latencies in the two tasks. Specifically, we measured participants' processing speed in non-linguistic tasks, computed how strongly processing speed correlated with lexical decision and picture naming latencies, and examined how strongly the correlation between the

---

[1] Note that the structure and contents of the mental lexicon is subject to extensive debate among cognitive scientists (see Taft, 2015, for a recent review), ranging from views where the mental representations accessed are "linguistically sophisticated" (Jackendoff, 2002) to views where the mental lexicon as a central storage is completely abandoned (Elman, 2011). Future research is needed to delineate between the various accounts. For the present study, we remain agnostic about the nature of the mental lexicon and use the term as a shorthand for stored word form information, which – without a doubt – language users must access to carry out picture naming and lexical decision tasks.

latencies in the two linguistic tasks was mediated by processing speed. No earlier research appears to have assessed the contribution of processing speed to performance in picture naming or lexical decision in young adults, but there is some research (predominantly on picture naming) on other age groups, which we briefly review.

In the cognitive aging literature, there is some discussion of the relationship between age-related general slowing and decrements in naming ability. General processing speed declines with age (e.g., Eckert, Keren, Robert, Calhoun, & Harris, 2010; Salthouse, 1996; Salthouse & Ferrer-Caja, 2003). The ability to access and retrieve words fast and accurately from the mental lexicon also declines with age (e.g., Mortensen, Meyer, & Humphreys, 2006; Myerson, Hale, Chen, & Lawrence, 1997), though in many studies of picture naming age effects on naming speed are small or absent (Belke & Meyer, 2007; Gordon & Kurzczek, 2014; Mortensen, Meyer, & Humphreys, 2008; Rizio, Moyer, & Diaz, 2017; Valente & Laganaro, 2015). It is very unlikely that the complex patterns of results observed in aging studies (for instance interactions of age with priming, frequency or neighborhood density effects, e.g., Burke & Shafto, 2004; Burke et al., 1991; Gordon & Kurzczek, 2014) are entirely due to age-related differences in processing speed, but general processing speed differences may contribute to these patterns. Two studies have directly assessed the relationship between general processing speed and picture naming speed in older adults. Soble et al. (2016) conducted a study with a sample (N=60) of healthy and neuro-cognitively impaired older participants and found that 26% of the variance in picture naming latency was explained by processing speed (measured as PSI, the processing speed index of the Wechsler Adult Intelligence Scale, 4th edition, WAIS-IV; Wechsler, 2008). Similarly, Rogalski, Peelle, and Reilly (2011) tested a sample of young adults (N=20, aged 18-26 years) and a sample of older adults (N=23, aged 54-81 years) in a picture naming task and two processing speed tasks (Trail Maker, TMT-A and TMT-B; Reitan & Wolfson, 1985). Consistent with the results obtained by Soble et al. (2016), they found that approximately 25% of the variance in picture naming latencies was accounted for by processing speed. Verhaegen and Poncelet (2013) reported that both general processing speed and naming speed declined with age, but did not quantify the relationship between the two skills.

Similarly, studies with children have found correlations between indicators of processing speed and performance in the Rapid Automatized Naming (RAN) task (e.g. Cutting & Denckla, 2001; Georgiou, Parrila, & Papadopoulos, 2016; Papadopoulos, Spanoudis, & Georgiou, 2016; Powell, Stainthorp, Stuart, Garwood, & Quinlan, 2007; Shaul & Nevo, 2015). In the RAN task, participants see a sheet of paper with, for instance, six rows of five objects, letters, or numbers, each. Each stimulus appears several times on the sheet. The task is to name as many of the stimuli as possible within a minute (Norton & Wolf, 2012). The participants' scores therefore do not only measure how fast participants name individual items, but also how quickly they can move their attention from item to item (Araújo et al., 2011; Gordon & Hoedemaker, 2016).

In sum, while studies of aging and development suggest that there may also be a relationship between processing speed and picture naming speed in young adults, no direct evidence about the strength of this link appears to be available. Since many studies have reported moderate correlations between processing speed and general intelligence (e.g., Sheppard & Vernon, 2008), we also included a test on non-verbal intelligence to separate the effects of non-verbal intelligence and processing speed from performance in the linguistic tasks.

Finally, earlier studies have shown that participants' vocabulary size predicted performance in picture naming and lexical decision tasks, with larger vocabularies being associated with faster and/or more accurate responses (see Brysbaert et al., 2016; Diependaele et al., 2013; Mainz et al., 2018). In addition, some studies have reported weaker word frequency effects in participants with larger, compared to smaller, vocabularies (Brysbaert, Lagrou, & Stevens, 2017; Diependaele et al., 2013; Yap, Tse, Balota, 2009). The origins of these relationships are not well understood. Effects of vocabulary size on lexical access speed may be due to structural differences, in that lexical representations in individuals with larger vocabularies are more robust or distinct, enabling faster processing, as compared to individuals with smaller vocabularies (Diependaele et al., 2013). That is, the interaction between word frequency and lexical access has been argued to result from differences in exposure to language (Brysbaert et al., 2016; Monaghan, Chang, Welbourne, & Brysbaert, 2017), where increased exposure is associated with an increase in efficiency of accessing lexical representations across the entire frequency range (Monaghan et al., 2017). Consequently, the lexicon of individuals with limited language exposure and therefore weaker word knowledge is hypothesized to show a stronger difference in processing efficiency between low- and high-frequency words due to less entrenched representations (see also Yap et al., 2009). Alternatively, cognitive skills, including processing speed, that allow individuals to acquire large lexica may also facilitate fast responding in laboratory tasks (see Monaghan et al., 2017). Note that the present study did not address the precise role of vocabulary size in lexical access. We included a test measuring participants' vocabulary size to be able to separate the effects of vocabulary size from any effects of processing speed on lexical access.

Measuring processing speed, non-verbal intelligence, and vocabulary size

Processing speed, sometimes called information-processing speed, speed of information processing, or general cognitive speed, is measured in elementary tasks, such as pressing a button as soon as a stimulus appears, where performance is unlikely to be strongly affected by different response strategies or prior experience with the task. However, even such simple tasks involve complex cognitive processes, including visual and conceptual processes, decision processes and sustained attention to the task. Consequently, there is much debate in the literature about the best way of measuring processing speed, whether processing speed is a unitary construct or whether several speed factors should be postulated, and how processing speed is related to other components of the cognitive system, in particular the attentional system and non-verbal intelligence (e.g., Cepeda, Blackwell, & Munakata, 2013; Jensen, 2006; Schubert, Hagemann, Voss, Schankin, & Bergmann, 2013; Schubert, Hagemann, & Frischkorn, 2017).

These controversial issues were not addressed in the present research. Our aim was to assess the hypothesis that participants might differ in their ability to carry out speeded laboratory tasks and to estimate the impact of this speed factor on the performance in the linguistic tasks. To this end, we used a battery of well-established short and simple processing speed tasks and two newly developed tasks and combined each participant's scores on these tasks into a single speed factor. This factor captured the shared variance across the tasks and served as our operational definition of processing speed. The well-established tasks were the digit-symbol substitution task (DSS) from the WAIS-IV

(Wechsler, 2008), the letter comparison task proposed by Earles and Salthouse (1995; Salthouse, 1996), and visual simple and choice reaction time tasks (e.g., Deary, Liewald, & Nissan, 2011; see Cepeda et al., 2013, for discussion). All of these tasks use visual stimuli. Since the present research concerns the processing of spoken words, we added a simple and a choice reaction time task using auditory stimuli. A methodological aim of the study was to test this battery of processing speed tasks, that is, to assess how well it could be used in laboratory and classroom settings, how reliable the retest scores would be, and how strongly they would correlate with each other. This should lead to practical recommendations about the assessment of processing speed in future research.

Non-verbal intelligence was assessed using a short version of Raven's Advanced Progressive Matrices (Raven et al., 1998). Vocabulary size was measured using the Peabody picture vocabulary test (PPVT; Dunn & Dunn, 1997).

Overview of the study

To recap, the present project addressed the following two research questions: First, to which extent are lexical access processes in word production and word comprehension related? Second, how strongly is lexical access in word-level processing moderated by domain-general non-verbal processing speed? To address these questions, young adult participants with diverse educational backgrounds carried out a picture naming and an auditory lexical decision task, as well as a battery of different processing speed tasks. To be able to separate the effects of processing speed, non-verbal intelligence, and vocabulary size on lexical access, we included two standard tests assessing non-verbal intelligence and vocabulary size, respectively.

Method

Participants

In total, 148 participants were tested. Eighty participants were students at Radboud University Nijmegen (RU students hereafter; 17 male, mean age of 22.2 years, range of 18 to 28 years), and 68 participants attended a vocational college in Doetinchem, a neighbouring city (VC students, 36 male, mean age of 19.9 years, range of 18 to 24 years). All students gave written informed consent to take part in the study and were paid 20 euro for participating. Permission to conduct the study was given by the Ethics Board of the Social Sciences Faculty of Radboud University.

Data from fifteen participants were excluded. They either did not complete all tasks (six VC students), or failed to follow the instructions for one or more tasks (one RU student), or had a history of neurological and/or developmental problems (one RU student). Picture naming data from four additional participants (all RU students) were lost due to technical issues. One RU student was excluded because of a strong yes-bias for the lexical decision task (100% accuracy for words, 20% for non-words) and two VC students were removed due to poor performance on the picture naming task (one with accuracy below 50%, one with average naming latencies of over 2000 ms). This left data from 133 participants (73 RU and 60 VC students) for the analyses.

Apparatus

RU students were tested on a Panasonic laptop with a 14-inch screen, running Windows 7. Auditory stimuli were presented using HD 201 Sennheiser headphones. VC students were tested using Hewlett Packard ‘ProBook’ laptops with 15-inch screens, also running Windows 7. Auditory stimuli were presented via Panasonic RP-HT030 headphones. In all experiments, visual stimuli were presented in black ink against a white background. Recordings from both participant groups were made using external Sennheiser microphones attached to the laptops. Except for the DSS task, which was administered using pen and paper, all tasks were implemented in Presentation© (version 16.5, www.neurobs.com). Hand-held button boxes with two buttons for left and right-hand thumbs were used for the reaction time (RT) tasks.

Procedure

The participants were tested in two sessions of about 65 and 45 minutes, respectively, separated by one week’s time. RU students were tested individually. VC students were tested in groups of 3 to 19 students in a quiet classroom setting on all tasks except picture naming, which was administered individually as the last test in the sequence. As the VC students were tested at their school in a different city than Nijmegen, individual testing was not possible due to logistic reasons. The order of the other tasks was identical to that in the RU students[2], as is common practice in individual differences studies. The motivation behind a fixed order of tasks is to minimize potential influences of the test procedure on participants' test performance. Table 1 shows the sequence of tests in each session. There were short breaks between the tests. The processing speed tasks were administered twice in separate sessions to assess their test-retest reliability. The materials, design, and procedure of the individual tests are described below.

Table 1. Order of tests in Sessions 1 and 2.

Session 1	*Session 2*
1. A-SRT	1. A-SRT
2. A-CRT	2. A-CRT
3. PPVT	3. Picture naming (8. in VC students)
4. V-SRT	4. V-SRT
5. V-CRT	5. V-CRT
6. Raven	6. LDT
7. Letter Comparison	7. Letter Comparison
8. DSS	8. DSS

Picture naming. The test materials consisted of 40 photographs of common objects, taken from de Groot et al. (2016) or retrieved online via a search engine (see Appendix A). The Dutch object names varied substantially in lexical frequency[3] (average ZipfF = 3.81, *SD* = 0.96, range = 1.85-5.64; as retrieved from the Subtlex Corpus, Keuleers et al., 2010). At the same time, prevalence ratings (Keuleers et al., 2015) suggested that all words were likely to be known by all participants (average prevalence 99.6%, *SD* = 0.4, range 97.7-100). The average number of phonological neighbors (sum of additions, substitutions, deletions of segments) of the object names was 4.33 (*SD* = 3.5, range = 0-16; as retrieved from Clearpond, Marian et al., 2012). Four additional photographs (ice skate, ashtray, lipstick, and pillow) were used as practice trials. All pictures were scaled to 300 x 300 pixels.

---

[2] RU students additionally carried out an eye-tracked pro-saccade task (Roberts, Hager, & Heron, 1994) as the first experiment of Session 1. Including eye-tracker calibration the task took around five minutes. This task was not administered to VC students; the results are therefore not discussed here.

[3] As suggested by van Heuven et al. (2014), we used Zipf frequencies, which were operationalized as: log10(frequency per million words) + 3.

Twenty participants (six male; mean age 22.4, range 18–33 years), native speakers of Dutch who did not take part in the main experiment, provided picture familiarity and visual complexity ratings in a study carried out at the Max Planck Institute for Psycholinguistics. The rating study was programmed in Presentation© and ran on an experimental laptop. The participants were presented with 80 photographs (order was randomized for each participant), one at a time, and were asked to rate each photograph's familiarity and then its visual complexity using a five-point scale (1 = unfamiliar/low visual complexity, 5 very familiar/high visual complexity). Participants could take a break after half of the trials. Forty of the 80 photographs were used in the present experiment. The mean familiarity rating score for those items was 3.8 (SD = 0.7, range = 2.55-5) and their mean visual complexity score was 2.8 (SD = 0.7, range = 1.6-3.95), suggesting about average visual complexity. These scores confirmed the overall suitability of the photographs for the picture naming task as participants were highly familiar with them and visual complexity was about average.

The task began with the presentation of four additional practice items. The order of the experimental items was randomized and different for each participant. On each trial, participants first saw a fixation cross in the center of the screen, which was shown for 800 ms, followed by the presentation of a blank screen for 200 ms. Then the target picture was shown for three seconds. After an inter-trial interval of one second, the next trial began. Participants were instructed to name the pictures as quickly as possible. Their utterances were recorded at a sampling rate of 44100 Hz, 16-bit resolution. Naming accuracy and onset latencies were coded offline using Praat software (Boersma, 2001).

Auditory lexical decision. Sixty Dutch words were selected from the Subtlex database (Keuleers et al., 2010). The words were not used as target names in the picture naming task, but were matched in frequency (average ZipfF = 3.65, *SD* = 0.85, range = 2.04-5.66), prevalence (average prevalence = 99.6, *SD* = 0.5, range = 97.3-100), and phonological neighborhood density (average = 2.8, *SD* = 3, range = 0-12). For each word, a matched pseudoword was created using Wuggy (Keuleers & Brysbaert, 2010) by substituting two phonemes without violating any phonotactic constraints of Dutch (see Appendix B). One additional word and two pseudowords were used on practice trials.

A female native speaker of Dutch, who had received professional speech training, produced first the words and then the pseudowords at a normal speech rate with neutral intonation. Her speech was recorded using a Sennheiser microphone sampling at a frequency of 44100 Hz (16-bit resolution). Audacity® software, version 2.5 (Audacity Team, 2014) and Praat (Boersma, 2001) were used to create speech files for the individual items. The average stimulus length was 655 ms (*SD* = 106, range = 451-1018) for words and 670 ms (*SD* = 124, range = 440-1010) for pseudowords.

The participants' task began with the presentation of the practice items. Experimental items followed, presented in a different random order for each participant. At the beginning of each trial, a fixation cross was presented in the center of the screen for 300 ms. Then one of the stimuli was presented. Participants had been instructed to listen carefully to each stimulus and decide as quickly as possibly whether or not it was an existing Dutch word. They pressed the right-hand button to give a "word" response and the left-hand button to give a "not a word" response. The response terminated the trial. After one second, the next trial began. The RT was the time interval between the spoken word onset and the button press.

Peabody picture vocabulary test (PPVT). The participants' receptive vocabulary size was assessed with a digitized version of the Dutch PPVT (Dunn & Dunn, 1997; Dutch translation by Schlichting, 2005). On each trial, participants first previewed four numbered line drawings on their screen. When they were ready, they pressed the Return key on their keyboard to hear the probe. They had to indicate which of the pictures best corresponded to the meaning of the spoken word by typing the corresponding number (1, 2, 3, or 4). Following the standard protocol for the test, items were presented in blocks of 12 items, with blocks increasing in difficulty. The starting level was 13, the best level participants could attain was 17. The test ended when a participant made nine or more errors within one block. Participants took on average 12 minutes to complete the test (range 8 to 15 minutes). The participants' score was their raw score, that is, the serial number of their last item minus the number of errors made during the test. Raw scores rather than percentile ranks based on Dutch norms (Schlichting, 2007) were used because several participants' scores fell below the normed range.

Advanced progressive matrices (APM). To assess non-verbal intelligence, a computerized version of Ravens' Advanced Progressive Matrices was used (Raven, Raven, & Court, 1998). On each trial, participants saw a panel of eight geometrical figures, with the space for a ninth figure left blank. From a set of eight candidates shown in the bottom section of the screen, they had to select the figure that completed the sequence. They indicated their choice by clicking on the chosen item with a mouse. Participants could skip items by clicking on button labelled "Skip"; these items were shown again at the end of the test. When they did not know the answer to a skipped item, participants could click on an "I don't know" button. There were 6 practice items and 36 test items, increasing in difficulty. Participants had 20 minutes to complete the test. Throughout the test, a clock in the right top corner of the screen showed the time remaining. A participant's score was their number of correct responses.

Digit-symbol substitution (DSS). This was a paper and pencil test, taken from the WAIS-IV (Wechsler, 2008). Participants received a worksheet showing a key of pairings of the digits 1 to 9 with arbitrary symbols (e.g., 1 = –; 2 = ⊥). Below the key, there were seven rows of 20 digits each. The participants' task was to write the corresponding symbol below each digit. The first seven digits in the first row were untimed practice items. After practice, participants had 90 seconds to complete the task. The score was the number of correct substitutions.

Letter comparison. This task was based on a paper-and-pencil task described in Earles and Salthouse (1995) and Salthouse (1996). We used the digitized version developed and described by Huettig and Janse (2016). On each trial, participants saw two letter strings (all consonants), one centered in the top half and one centered in the bottom half of the screen. The letter strings were identical or differed by one letter. Participants had to decide as quickly and as accurately as possible whether the two letter strings were the same or different by pressing the right-hand button ("same") or the left-hand button ("different") on their button box. The first test block featured three-letter strings (e.g. TZF) and the second block six-letter strings (e.g. RNHKTG). Letters were presented in a large black font (Arial 60) against a white background. Within each test block, 12 pairs were identical and 12 were different. A practice block of three identical and three different pairs preceded the first block.

Each trial started with the presentation of a fixation cross, which stayed on the screen for 500 ms. After a blank interval of 100 ms, the two letter strings were presented and stayed on the screen until the participant responded. The next trial started after an inter-trial time of one second. Participants' score was their average reaction time calculated based on trials that were responded to correctly.

Visual simple reaction time task (V-SRT). On each trial, participants first saw a fixation cross in the center of the screen. After an interval varying randomly between one and three seconds, it was replaced by a line drawing of a triangle (200 x 200 pixels, black contours). Participants were instructed to press the right-hand button on their button box as soon as the triangle appeared. The response terminated the trial. After a blank interval of one second, the next trial began. The task consisted of eight practice trials followed by 20 test trials. Participants' score was their average reaction time on test trials.

Visual choice reaction time task (V-CRT). As in the V-SRT, trials began with the presentation of a fixation cross for one to three seconds. Then a line drawing of either a star or a circle (black contours, 200 x 200 pixels) appeared on the screen (cf. Cepeda et al., 2013). Participants pressed the left-hand button on their button box as fast as possible upon appearance of a star, and the right-hand button upon appearance of the circle. The star and circle appeared equally often in random order. There were 16 practice trials followed by 40 test trials. Participants' score was their average reaction time calculated based on trials that were responded to correctly.

Auditory simple reaction time task (A-SRT). The auditory reaction time tasks were designed to be as similar as possible to the visual reaction time tasks. Each trial of the A-SRT began with the presentation of the fixation cross, which now remained in view until participants responded. One to three seconds after trial onset (varying interval), a sine tone (550 Hz, 400 ms) was played. Participants pressed the right-hand button on their button box as soon as they heard the tone. This terminated the trial. After one second, the next trial began. The test consisted of a practice block of eight trials followed by 20 test trials. As in the V-SRT, participants' score was their average reaction time on test trials.

Auditory choice reaction time task (A-CRT). As in the A-SRT, the trial began with the presentation of the fixation cross. After a varied interval between one and three seconds, either a high (800 Hz, 400 ms) or a lower sine tone (300 Hz, 400 ms) was played. Participants responded as quickly as possible by pressing the right-hand button when they heard the high tone, and the left-hand button, when they heard the low tone. There was a practice block of 16 trials, followed by 40 test trials and their score was their average reaction time calculated based on trials that were responded to correctly.

Results

Processing speed, non-verbal intelligence and vocabulary size tasks

The results obtained for the six processing speed tasks, the vocabulary test (PPVT) and the test of non-verbal intelligence (APM) are summarized in Table 2. The error rates in the auditory and visual response time tasks and in the letter comparison task were low. Inspection of the data yielded no evidence for a speed-accuracy tradeoff. Error trials were eliminated from the data set, as were any outliers, defined as RTs deviating from a participant's mean by more than 2.5 SD. The test-retest reliability of the test scores was established by conducting two-tailed Pearson correlations between participants' scores on Session 1 and Session 2. As the table shows, retest reliability was satisfactory for all tasks and somewhat higher for the more complex ones than for the simpler ones.

Table 3 shows the correlations between the scores in all non-linguistic tasks, for the first session only. For the following analyses, the scores on the processing speed tasks for Session 1 were combined into a single measure by creating a processing speed factor (PS_Fac). To that end, regression-based factor scores were calculated for each participant using the principal component analysis (PCA) method in SPSS (no rotation). Bartlett's Test of Sphericity, $\chi^2(15) = 288.38$, $p < .001$, confirmed that the strengths of the correlations permitted a PCA. The Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis KMO = .81 ('great' according to Kaiser, 1974). The outcome of the PCA revealed that one component had an eigenvalue above Kaiser's criterion of 1 and explained 54% of the variance. All six processing speed tasks loaded highly on this factor, with values above .50 (see Table 2).

Table 2. Results for processing speed, non-verbal intelligence and vocabulary size tasks.

Task	Score	Error Rate (%)	Outliers (%)	Retest Reliability	Factor loading
A-SRT 1	251 (40)	0 (0)	2.7 (2.6)	0.69 (<.001)	0.76
A-SRT 2	247 (34)	0 (0)	2.9 (2.5)	-	-
A-CRT 1	405 (66)	2.7 (2.7)	2.1 (1.6)	0.72 (<.001)	0.84
A-CRT 2	396 (71)	3.7 (3.9)	2.5 (1.6)	-	-
V-SRT 1	241 (32)	0 (0)	3.2 (2.6)	0.61 (<.001)	0.78
V-SRT 2	244 (37)	0 (0)	3.5 (2.6)	-	-
V-CRT 1	403 (59)	3.1 (3.2)	2.3 (1.6)	0.68 (<.001)	0.79
V-CRT 2	400 (56)	3.3 (3.6)	2.2 (1.5)	-	-
Letter 1	1087 (253)	6.1 (4.1)	1.9 (1.4)	0.81 (<.001)	0.68
Letter 2	974 (212)	6.3 (4.9)	1.7 (1.4)	-	-
DSS 1	64 (12)	-	-	0.84 (<.001)	-0.55
DSS 2	72 (11)	-	-	-	-
PPVT	170 (12)	-	-	-	-
APM	19 (7)	-	-	-	-

Note: Standard deviations provided in brackets for Score, Error Rate, and Outliers. Retest reliability operationalized as two-tailed Pearson Correlation Coefficient, p-values provided in brackets. A-S/CRT, V-S/CRT, and Letter scores in ms; DSS and APM scores reflect number of correct responses; PPVT scores reflect number of last test item minus number of errors.

Table 3. Correlations between processing speed, non-verbal intelligence and vocabulary size tasks.

	A-CRT	V-SRT	V-CRT	Letter	DSS	PPVT	APM
A-SRT	.66**	.54**	.43**	.30**	-.33**	-.45**	-.29**
A-CRT		.60**	.57**	.42**	-.33**	-.44**	-.28**
V-SRT			.57**	.37**	-.25**	-.31**	-.15
V-CRT				.54**	-.30**	-.26**	-.14
Letter					-.42**	-.22*	-.23**
DSS						.24**	.37**
PPVT							.50**

Note: *p < .05, **p <.01. A/V-S/CRT = Auditory/visual simple/choice reaction time task. Letter = Letter comparison task. DSS = Digit-symbol substitution task. PPVT = Peabody picture vocabulary test. APM = Advanced progressive matrices.

Linguistic tasks

In the analyses of the linguistic tasks, we first examined how well the participants' response speed and accuracy in each of the two tasks, picture naming and lexical decision, was predicted by their vocabulary size, non-verbal intelligence, and processing speed and by the Zipf frequency of the items, and then examined how strongly the participants' performance scores in the two linguistic tasks were related.

Picture naming. Five items were removed from the picture naming data set because the error rates exceeded 50%. Error trials were removed from the latency analyses, as were trials with latencies deviating by more than 2.5 SD from the participant's mean (see Table 4).

Table 4. Results for the linguistic tasks.

Task measure	Mean (ms)	Error rate (%)	Outlier rate (%)
Picture naming	789 (120)	14.9 (8.9)	2.4 (1.9)
LD words	880 (91)	4.8 (3.8)	0.9 (1.2)
LD pseudowords	1088 (169)	12.3 (9.7)	2.2 (1.4)
LD d-prime	3.1 (0.6)	-	-

Note: Standard deviations in brackets.

Naming latencies were analyzed by means of a linear mixed effects model using the lme4 package (Bates, Mächler, & Bolker, 2015) in R (R core team, 2012). The model contained four continuous predictors: Zipf frequency, processing speed (PS-Fac), non-verbal intelligence (APM), and vocabulary size (PPVT). All predictors were centered and scaled. Participants and Items were included as crossed random factors. For participant, the random slope for Zipf frequency was added. Predictors that did not reliably contribute to model fit were dropped. Models were compared using likelihood ratio tests. Starting from the full model, dropping APM and PPVT did not result in worse model fit, $\chi^2(1) = 0.79$, $p = .37$ and $\chi^2(1) = 0.08$, $p = .78$, respectively. Removing Zipf frequency had a marginal effect: $\chi^2(1) = 2.89$, $p = .09$. As was to be expected, higher Zipf frequency was associated with faster onset latencies. Finally, removing PS-Fac led to a significant decrease in model fit; $\chi^2(1) = 17.58$, $p < .001$. The best-fitting model therefore only included the predictor PS-Fac ($\beta = 40.99$, SE = 9.48, t = 4.32).

Naming accuracy was analyzed using logit mixed models (Jaeger, 2008). The model structure and procedure to remove non-contributing predictors were the same as in the onset latency analyses, except that the random slope for Zipf frequency was not included as this led to failure in model convergence. While removing PPVT did not result in worse model fit ($\chi^2(1) = 1.06$, $p = .31$), dropping any of the three other factors did, APM: $\chi^2(1) = 8.57$, $p < .01$; PS-Fac $\chi^2(1) = 4.90$, $p < .05$; Zipf frequency: $\chi^2(1) = 5.11$, $p < .05$ (see Table 5 for the model estimates for the three factors).

In sum, the frequency of the picture names had a weak effect on the picture naming latencies and a more robust effect on accuracy. Processing speed predicted both speed and accuracy of picture naming, and non-verbal intelligence predicted picture naming accuracy but not speed.

Table 5. Model estimates for picture naming accuracy and onset latency.

Fixed effect	Estimated Coefficient (β)	Standard Error (SE)	t/z-value	p-value
Onset latency				
Speed	40.99	9.48	4.32	<.001
Accuracy				
Frequency	0.61	0.26	2.32	<.05
Speed	-0.16	0.07	-2.25	<.05
APM	0.28	0.07	3.75	<.001

Lexical decision. The results obtained in the lexical decision task are summarized in Table 4. Participants were faster and more accurate in their responses to words than to pseudowords. As word frequency was only relevant for responses to words, latencies for word and pseudo-word responses were analyzed separately. Error trials were excluded from the latency analyses, as were responses deviating more than 2.5 SD from a participant's mean RT. For word responses, a linear mixed effects model was run with the same structure as used for picture naming RT. Removing PPVT did not result in worse model fit ($\chi^2(1) = 0.01$, $p = .90$). Dropping any of the three other factors did affect model fit, Zipf frequency: $\chi^2(1) = 9.86$, $p < .01$; PS-Fac $\chi^2(1) = 44.90$, $p < .001$; APM: $\chi^2(1) = 7.04$, $p < .01$ (see Table 6 for the results of the best fitting model). Surprisingly, higher APM scores were associated with slower responses[4].

The linear mixed effects model for pseudoword RT did not include the fixed factor frequency, nor the random slope for word frequency for Participants. Removing PPVT did not result in worse model fit ($\chi^2(1) = 1.34$, $p = .25$), nor did removing APM, $\chi^2(1) = 2.62$, $p = .11$. Only removing PS-Fac

---

[4] Based on a reviewer suggestion, we also ran a model that included 'word length' as a continuous predictor, the interaction of it with 'Zipf frequency', and 'word length' as random slope by 'Participant'. The best fitting model yielded significant effects of 'word length' (β = 31.62, SE = 9.61, t = 3.29), 'Zipf frequency' (β = -38.85, SE = 9.6, t = -4.05), 'PS-Fac' (β = 50.78, SE = 6.96, t = 7.29), and APM (β = 18.9, SE = 7.01, t = 2.7). Thus, word length contributed significantly to explaining variance in the dependent variable. Importantly, the contributions of the other predictor variables were unaffected by 'word length' being in the model.

significantly decreased model fit: $\chi^2(1) = 22.49$, $p < .001$. Thus, the best-fitting model only included the predictor PS-Fac (β = 63.24, SE = 12.85, t = 4.92).

For both words and pseudowords, faster responses were related to fast general processing speed. High general intelligence only affected the speed of word, but not pseudoword responses. Finally, word frequency had the expected effect such that high frequency words were responded to faster than low frequency words.

Table 6. Model estimates for lexical decision word and pseudoword response times.

Fixed effect	Estimated Coefficient (β)	Standard Error	t/z-value	p-value
Word response times				
Frequency	-32.91	10.20	-3.23	<.01
Speed	50.80	6.96	7.30	<.001
APM	18.82	7.00	2.69	<.01
Pseudoword response times				
Speed	63.24	12.85	4.92	<.001
Accuracy (d')				
PPVT	0.11	0.05	2.11	<.01
APM	0.17	0.05	3.28	<.01

Participants made more errors on pseudoword than on word trials, indicating that they were more likely to provide word than pseudoword responses. To take this bias into account in computing accuracy scores, we used d', calculated as d' = Z(proportion correctly identified words) - Z(proportion incorrectly identified non-words) for all 133 participants. As d' cannot be calculated for hit or false alarm rates of 0 or 1, these values were replaced with 0.5/n and (n-0.5)/n respectively (MacMillan & Kaplan, 1985; Stanislaw & Todorov, 1999) The d'-values were analyzed using multiple linear regression, with the predictors PS-Fac, APM, and PPVT. PS-Fac did not contribute significantly to the model, as dropping this factor from the model did not result in worse fit: $\chi^2(1) = 0.01$, $p > .9$. Dropping either of the remaining factors did, PPVT: $\chi^2(1) = 4.44$, $p < .05$; APM: $\chi^2(1) = 10.73$, $p < .01$.The results revealed that participants with larger vocabularies and higher intelligence scores had higher d' scores (see Table 6).

Relationship between picture naming and lexical decision latencies. Finally, we investigated the relationship between picture naming and lexical decision latencies for words. The correlation of the participants' correct response latencies was r = .40 (p < .001; see Figure 1, also for other inter-measure correlations). Controlling for APM barely changed the correlation (r = .42), nor did controlling for PPVT (r = .39) or both variables simultaneously (r = .41). However, controlling for PS-Fac reduced the correlation to r = .32 (p <.001). In other words, general processing speed accounted for part of the shared variance in picture naming onset latencies and lexical decision response times. The correlation between the proportion of correct responses in the two tasks was not significant, r = .15 (p = .08). Therefore, no partial correlations were computed.

Discussion

In this study, young adults, who attended university or vocational training courses, completed a picture naming and a lexical decision tasks as well as a battery of processing speed tasks, a test of non-verbal intelligence, and a vocabulary test. Picture naming and lexical decision are widely considered good tools to tap into lexical access. The main goal of the study was to assess the extent to which lexical access processes are shared across production and comprehension, that is, how strongly performance on both tasks was correlated. Moreover, given the nature of these tasks (i.e., that they both involve a speeded response), a second goal was to assess the extent to which the two tasks rely on domain-general abilities (specifically, processing speed). We start by summarizing the results for the picture naming and lexical decisions tasks, respectively. In doing so, we discuss how strongly performance on each of the tasks was influenced by processing speed before addressing the main research question. Finally, we provide practical recommendations pertaining to the assessment of processing speed.

Picture naming

For picture naming, we found effects of the control variables Zipf frequency and APM on the error rates, but no effects of any of the control variables on the naming latencies. Thus, more errors were made for less frequent items and by participants with lower APM scores (interpreted as lower non-verbal intelligence). The only marginal word frequency effect on the picture naming latencies may come as a surprise, as many earlier studies have reported that picture naming latencies depend on name frequency (Jescheniak & Levelt, 1994; Oldfield & Wingfield, 1965). It is possible that the frequency range in the present study was not sufficient to detect a frequency effect (Brysbaert & Cortese, 2011). In the present study, frequency was a continuous variable, whereas earlier studies have typically contrasted sets of pictures with high frequency versus low frequency names, which may have facilitated detecting frequency effects. It is also possible that a name frequency effect was concealed by effects of other variables that were not controlled for, for instance the complexity of the images or the familiarity of the objects. However, at least for the present materials, this option appears to be unlikely as our rating studies ensured high and homogenous picture familiarity as well as average visual complexity across the entire set.

Performance on the picture naming task was influenced by general processing speed: Participants with high processing speed showed faster and more accurate responses to the pictures than participants with lower processing speed. Given the speeded nature of the picture naming task, one might not find this result particularly surprising. Nevertheless, the present study is, to the best of our knowledge, the first to demonstrate that this relationship between naming performance and processing speed indeed exists in a group of healthy young adults. This result raises a number of questions for further research. One may for instance ask whether similar effects of processing speed will be seen in other language production tasks or in other groups of speakers. More interestingly, one may ask how the correlation between processing speed and picture naming performance arose. Is there a causal link, such that speakers endowed with high processing speed can complete the complex processes underlying picture naming faster than speakers with lower processing speed? If so, are all components of the naming process equally affected by variations in processing speed, in proportion to their durations, or are some components affected more strongly than others (e.g., Laganaro, Valente & Perret, 2012)? These questions can be addressed in further research examining whether processing speed scores interact with variables affecting specific components of the naming process, such as conceptual preparation, lexical selection, or articulation.

Alternatively, there might not be a direct, but rather a mediated relationship between processing speed and naming. Processing speed has been shown to correlate with attention and executive control skills (Salthouse & Meinz, 1995; Vernon, 1987). Therefore, the correlation between picture naming latencies and processing speed may be mediated by variation in one or more of these skills, for instance sustained attention skills. This important issue, of how the relationship between picture naming and processing speed arises, can be addressed in further studies where participants complete batteries of tasks, tapping not only processing speed, but also attention and executive control. The unique impact of different domain-specific skills on naming performance could then be determined.

Lexical decision

For the error rates in the lexical decision task, we found effects of the control variables non-verbal intelligence and vocabulary, but no effect of processing speed. For the lexical decision latencies, we found effects of word frequency, non-verbal intelligence, and processing speed. The frequency effect is consistent with a large body of earlier work showing that lexical decisions are made faster to high than to low frequency words (e.g., Ratcliff & McKoon, 1997; Goldinger, 1996, for review). The direction of the effect of non-verbal intelligence was unexpected, as higher APM scores were associated with slower rather than faster responses. This suggests that individuals with higher non-verbal intelligence might have been more careful when responding to the word stimuli. That is, participants with higher APM scores responded more slowly and made fewer errors than participants with lower APM scores.

The most important result is, again, an effect of processing speed. Participants with high processing speed responded faster to the words and pseudowords than participants with lower processing speed. Although one might once again have expected this outcome, the present study shows for the first time that response times in the lexical decision task are substantially influenced by individuals’ overall processing speed. As discussed for the picture naming task, further research could determine which of the processes involved in lexical decision are most strongly affected by variations in processing speed and whether there is a direct or a mediated link between processing speed and lexical decision speed.

Shared lexical access processes in speaking and listening

Turning to the main question, the relationship between picture naming and lexical decision, we obtained a correlation of r = .40 between the latencies in the two tasks. The strength of the correlation did not change when the impact of non-verbal intelligence and vocabulary size was controlled for, but dropped to r = .32 when processing speed was controlled for. Thus, only 10% of the variance was shared between the two tasks. Considering the nature of the two tasks, finding only a small portion of shared variance is not too surprising. Though both tasks are speeded, and both involve lexical access, picture naming uniquely involves visual and conceptual encoding processes for complex images and the preparation of articulatory responses, while lexical decision uniquely involves the acoustic and phonetic processing of the stimuli and metalinguistic decision processes. These input and output processes apparently involve skills that are not strongly correlated with each other or with the non-linguistic skills assessed here (see also Balota & Chumbley, 1984; Broos, Duyker & Hartsuiker, 2018).

Though most of the variance in the lexical decision and picture naming tasks was not shared, the small amount of shared variance should not be overlooked. Further research could seek to identify the underlying shared processes. These processes may be domain-general processes, such as inhibitory control or sustained attention, which were not assessed here. Alternatively, they may be more specific to lexical access, for instance concerning the ability to discriminate rapidly between co-activated lexical representations or concerning the strength of links between form and meaning representations (Burke et al., 1991; Britt et al., 2016; Dell et al., 1997; Gordon & Kurzcek, 2014). One way of exploring these and related hypotheses is to ask participants to complete batteries of production and comprehension tasks as well batteries of non-linguistic tasks and use statistical modelling to identify the underlying shared traits. Such an approach is occasionally taken in studies of language aptitude, which is often conceptualized as the ability to learn a second or further language (Andringa et al., 2012; Trapman et al., 2014; see also Hulstijn, 2018), but may also be very productive in research on individual differences in native language skills.

Assessment of processing speed

The processing speed battery run in the present study consisted of a visual and an auditory simple reaction time task, a visual and an auditory choice reaction time task, the digit-symbol substitution and the letter-comparison task. All tasks were easy to administer both in lab and classroom settings. The retest reliability for all tests, including the newly developed auditory processing speed tasks, was satisfactory. Reliability was slightly higher for the more complex than for the simpler tasks, possibly due to lower variability on the simpler tasks (cf. Soveri et al., 2018). All processing speed scores correlated significantly with each other, and in the principle component analysis only one significant factor was extracted. The simple and choice reaction time tasks had the highest loadings on this factor. The loadings of the two more complex tasks on this factor were lower (though still above .40), but these tasks did not load highly on a second factor. Thus, there was no evidence that the tests using auditory versus visual stimuli, or those using simple versus more complex tasks captured different abilities. Instead, the PCA pointed to the existence of a single underlying trait, processing speed.

Participants' processing speed scores correlated moderately with the scores obtained on the vocabulary and non-verbal intelligence tests. This pattern is consistent with results of earlier studies (e.g., Monaghan et al., 2017; Schubert et al., 2015, 2017). As explained in the Introduction, the causal relationships between these variables remain to be uncovered. For instance, high processing speed is likely to contribute to good performance in the non-verbal intelligence tests, and it is likely to support word learning and therefore to lead to high scores on vocabulary tests (Monaghan et al., 2017). Similarly, high scores on non-verbal general intelligence tests may indicate inferencing strategies that also support extracting the meaning of novel words from texts (e.g., Mainz, 2018). An important goal for further research into the causes of individual differences in linguistic skills is to unravel the complex relationships between non-verbal intelligence, processing speed, and vocabulary.

In sum, our data showed that the processing speed tasks used in the present study loaded more or less equally strong onto the processing speed factor and that all had acceptable retest reliability. Given these results, one might be tempted to conclude that future studies do not necessarily need to run a whole battery, and that one or two tasks suffice to capture participants' processing speed skills. Clearly, for most studies six tasks are not needed. However, in keeping with previous proposals, we recommend including at least three tasks—three that seem most different from one another at the surface to minimize the so-called task impurity problem (Miyake et al., 2000, for further discussion).

Conclusion

Lexical access in word production and word comprehension is often assessed using picture naming and auditory lexical decision, respectively. Results from these tasks have contributed substantially to formulating theories and models of word processing in speaking and listening. Here, we have shown that lexical access processes as measured by these tasks shared about 10% of common variance and that performance on both tasks was strongly influenced by individuals' general processing speed. Given this pattern of results and in particular the rather weak relationship between picture naming and lexical decision, we conclude that these tasks are not suited to *selectively* assess lexical access ability in an individual or a group. Instead, in order to obtain purer measures of lexical access, performance on these tasks must be considered together with performance in other cognitive tasks, including other tasks that tap into lexical access, tasks that tap into other language processing components, and tasks assessing domain-general skills involved in lexical access. It is therefore indispensable that studies of individual differences in lexical access routinely administer multiple word-level processing tasks as well as multiple non-linguistic processing speed tasks to be able to extract shared variance representing the skill at stake (e.g., Christopher et al., 2016). The present individual differences study nevertheless already suggests that the lexical access processes used in speaking are shared to only a limited extent with those used in listening.

References

Adelman, J. S., Sabatos-DeVito, M. G., Marquis, S. J., & Estes, Z. (2014). Individual differences in reading aloud: A mega-study, item effects, and some models. *Cognitive Psychology, 68,* 113-160.

Anderson, S. E., Chiu, E., Huette, S., & Spivey, M. J. (2011). On the temporal dynamics of language-mediated vision and vision-mediated language. *Acta Psychologica*, *137*(2), 181-189.

Andrews, S. (2012). Individual differences in skilled visual word recognition and reading. *Visual word recognition, 2,* 151-172.

Andringa, S., Olsthoorn, N., van Beuningen, C., Schoonen, R., & Hulstijn, J. (2012). Determinants of success in native and non-native listening comprehension: an individual differences approach. *Language Learning, 62*(Suppl. 2), 49-78.

Araújo, S., Inácio, F., Francisco, A., Faísca, L., Petersson, K. M., & Reis, A. (2011). Component processes subserving rapid automatized naming in dyslexic and non-dyslexic readers. *Dyslexia*, *17*(3), 242-255.

Audacity Team (2014). Audacity(R): Free Audio Editor and Recorder [Computer application]. Version 2.5.0 retrieved [date] from https://audacityteam.org/.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of word frequency in the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(3), 340-357.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Belke, E., & Meyer, A. S. (2007). Single and multiple object naming in healthy ageing. *Language and Cognitive Processes*, *22*(8), 1178-1211.

Bock, J. K. & Ferreira, V. S. (2014). Syntactically speaking. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds), *The Oxford Handbook of Language Production* (pp. 21-46). New York: Oxford University Press.

Boersma, P. P. G. (2002). Praat, a system for doing phonetics by computer (Version 5.1.19) [Computer program]. Available from http://www.praat.org/.

Britt, A. E., Ferrara, C., & Mirman, D. (2016). Distinct effects of lexical and semantic competition during picture naming in younger adults, older adults, and people with aphasia. *Frontiers in Psychology, 7,* 813.

Broos, W. P., Duyck, W., & Hartsuiker, R. J. (2018). Are higher-level processes delayed in second language word production? Evidence from picture naming and phoneme monitoring. *Language, Cognition and Neuroscience*, 1-16.

Brysbaert, M., Lagrou, E., & Stevens, M. (2017). Visual word recognition in a second language: A test of the lexical entrenchment hypothesis with lexical decision times. *Bilingualism: Language and Cognition*, *20*(3), 530-548.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). The impact of word prevalence on lexical decision times: Evidence from the Dutch Lexicon Project 2. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(3), 441-458

Brysbaert, M., & Cortese, M. J. (2010). Do the effects of subjective frequency and age of acquisition survive better word frequency norms?. *The Quarterly Journal of Experimental Psychology*, *64*(3), 545-559.

Burke, D. M., MacKay, D. G., Worthley, J. S., & Wade, E. (1991). On the tip of the tongue: What causes word finding failures in young and older adults?. *Journal of Memory and Language*, *30*(5), 542-579.

Burke, D. M., & Shafto, M. A. (2004). Aging and language production. *Current Directions in Psychological Science*, *13*(1), 21-24.

Cepeda, N. J., Blackwell, K. A., & Munakata, Y. (2013). Speed isn't everything: Complex processing speed measures mask individual differences and developmental changes in executive control. *Developmental Science*, *16*(2), 269-286.

Chater, N., McCauley, S. M., & Christiansen, M. H. (2016). Language as skill: Intertwining comprehension and production. *Journal of Memory and Language, 89*, 244-254.

Christopher, M. E., Keenan, J. M., Hulslander, J., DeFries, J. C., Miyake, A., Wadsworth, S. J., Willcutt, E., Pennington, B., & Olson, R. K. (2016). The genetic and environmental etiologies of the relations between cognitive skills and components of reading ability. *Journal of Experimental Psychology: General*, *145*(4), 451.

Cutting, L. E., & Denckla, M. B. (2001). The relationship of rapid serial naming and word reading in normally developing readers: An exploratory model. *Reading and Writing, 14*(7-8), 673-705.

Dąbrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. *Cognition, 178*, 222-235.

Deary, I. J., Liewald, D., & Nissan, J. (2011). A free, easy-to-use, computer-based simple and four-choice reaction time programme: the Deary-Liewald reaction time task. *Behavior Research Methods, 43*(1), 258-268.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104*(4), 801.

Diependaele, K., Lemhöfer, K., & Brysbaert, M. (2013). The word frequency effect in first and second language word recognition: A lexical entrenchment account. *The Quarterly Journal of Experimental Psychology, 66*(5), 843-863.

Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary. Test–Third Edition–Dutch Adaptation (L. Schlichting, Trans., 2005). Amsterdam, the Netherlands: Harcourt Test.

Earles, J. L., & Salthouse, T. A. (1995). Interrelations of age, health, and speed. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *50*(1), P33-P41.

Eckert, M. A., Keren, N. I., Robert, D. R., Calhoun, V. D., & Harris, K. C. (2010). Age-related changes in processing speed: unique contributions of cerebellar and prefrontal cortex. *Frontiers in Human Neuroscience*, *4*, 10.

Elman, J. L. (2011). Lexical knowledge without a lexicon?. *The Mental Lexicon 6* (1), 1–33.

Engelhardt, P. E., Nigg, J. T., & Ferreira, F. (2017). Executive function and intelligence in the resolution of temporary syntactic ambiguity: An individual differences investigation. *The Quarterly Journal of Experimental Psychology*, *70*, 1263-1281.

Georgiou, G. K., Parrila, R., & Papadopoulos, T. C. (2016). The anatomy of the RAN-reading relationship. *Reading and Writing*, *29*(9), 1793-1815.

Goldinger, S. D. (1996). Auditory lexical decision. *Language and Cognitive Processes*, *11*(6), 559-568.

Gordon, J. K., & Kurczek, J. C. (2014). The ageing neighbourhood: Phonological density in naming. *Language, Cognition and Neuroscience*, *29*(3), 326-344.

Gordon, P. C., & Hoedemaker, R. S. (2016). Effective scheduling of looking and talking during rapid automatized naming. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(5), 742.

De Groot, F., Koelewijn, T., Huettig, F., & Olivers, C. N. L. (2016). A stimulus set of words and pictures matched for visual and semantic similarity. *Journal of Cognitive Psychology, 28*(1), 1-15.

Hagoort, P. (Ed.). (in press). Human Language: From Genes and Brains to Behavior. Cambridge, MA: MIT Press.

Huettig, F., & Janse, E. (2016). Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. *Language, Cognition and Neuroscience, 31*(1), 80-93.

Hulstijn, J. H. (2018). An individual-differences framework for comparing non-native with native speakers: Perspectives from BLC Theory. *Language Learning,* Advanced online publication.

Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1-2), 101-144.

Jackendoff, R.S. (2002). *Foundations of Language: Brain, Meaning, Grammar, and Evolution.* Oxford University Press.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434-446.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Elsevier.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 824-843.

Kaiser, H. 1974. An index of factor simplicity. *Psychometrika* 39: 31–36.

Keuleers, E., Stevens, M., Mandera, P., & Brysbaert, M. (2015). Word knowledge in the crowd: Measuring vocabulary size and word prevalence in a massive online experiment. *The Quarterly Journal of Experimental Psychology, 68*(8), 1665-1692.

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods*, *42*(3), 627-633.

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643-650.

Kidd, E., Donnelly, S., & Christiansen, M. H. (2018). Individual differences in language acquisition and processing. *Trends in Cognitive Sciences, 22*(2), 154-169.

Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, *73*, 1-40.

Laganaro, M., Valente, A., & Perret, C. (2012). Time course of word production in fast and slow speakers: a high density ERP topographic study. *NeuroImage*, *59*(4), 3881-3888.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*(1), 1-38.

Litcofsky, K. A., Tanner, D., & van Hell, J. G. (2015). Effects of language experience, use, and cognitive functioning on bilingual word production and comprehension. *International Journal of Bilingualism, 20*(6), 666-683.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*(1), 1.

Macmillan, N. A., & Kaplan, H. L. (1985). Detection theory analysis of group data: estimating sensitivity from average hit and false-alarm rates. *Psychological Bulletin*, *98*(1), 185.

Mainz, N. (2018). Vocabulary knowledge and learning: Individual differences in adult native speakers. PhD Thesis, Radboud University Nijmegen, Nijmegen.

Mainz, N., Shao, Z., Brysbaert, M., & Meyer, A. S. (2017). Vocabulary Knowledge Predicts Lexical Processing: Evidence from a Group of Participants with Diverse Educational Backgrounds. *Frontiers in Psychology, 8*: 1164.

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PloS One*, *7*(8), e43230.

McQueen, J. M., & Meyer, A. S. (in press). Towards a comprehensive cognitive architecture for language use. In P. Hagoort (Ed.), *Human language: From Genes and Brain to Behavior*. Cambridge, MA: MIT Press.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100.

Monaghan, P., Chang, Y. N., Welbourne, S., & Brysbaert, M. (2017). Exploring the relations between word frequency, language exposure, and bilingualism in a computational model of reading. *Journal of Memory and Language*, *93,* 1-21.

Mortensen, L., Meyer, A. S., & Humphreys, G. W. (2006). Age-related effects on speech production: A review. *Language and Cognitive Processes*, *21*(1-3), 238-290.

Mortensen, L., Meyer, A. S., & Humphreys, G. W. (2008). Speech planning during multiple-object naming: Effects of ageing. *The Quarterly Journal of Experimental Psychology*, *61*(8), 1217-1238.

Myerson, J., Hale, S., Chen, J., & Lawrence, B. (1997). General lexical slowing and the semantic priming effect: The roles of age and ability. *Acta Psychologica*, *96*(1-2), 83-101.

Norton, E. S., & Wolf, M. (2012). Rapid automatized naming (RAN) and reading fluency: Implications for understanding and treatment of reading disabilities. *Annual Review of Psychology, 63*, 427-452.

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *The Quarterly Journal of Experimental Psychology*, *17*(4), 273-281.

Papadopoulos, T. C., Spanoudis, G. C., & Georgiou, G. K. (2016). How is RAN related to reading fluency? A comprehensive examination of the prominent theoretical accounts. *Frontiers in Psychology*, *7,* 1217.

Peramunage, D., Blumstein, S. E., Myers, E. B., Goldrick, M., & Baese-Berk, M. (2011). Phonological neighborhood effects in spoken word production: An fMRI study. *Journal of Cognitive Neuroscience*, *23*(3), 593-603.

Powell, D., Stainthorp, R., Stuart, M., Garwood, H., & Quinlan, P. (2007). An experimental comparison between rival theories of rapid automatized naming performance and its relationship to reading. *Journal of Experimental Child Psychology*, *98*(1), 46-68.

Presentation® software. Version 16.5, Neurobehavioral Systems, Inc., Berkeley, CA, www.neurobs.com.

R Development Core Team. (2012). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Open access available at: http://cran.rproject.org

Ratcliff, R., & McKoon, G. (1997). A counter model for implicit priming in perceptual word identification. *Psychological Review*, *104*, 319-343.

Raven, J., Raven, J. C., & Court, J. H. (1998). Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 4: *The Advanced Progressive Matrices*. San Antonio, TX: Harcourt Assessment.

Reitan, R. M., Wolfson, D. (1985). *The Halstead–Reitan Neuropsychological Test Battery: Therapy and clinical interpretation*. Tucson, AZ Neuropsychological Press

Rizio, A. A., Moyer, K. J., & Diaz, M. T. (2017). Neural evidence for phonologically based language production deficits in older adults: An fMRI investigation of age-related differences in picture-word interference. *Brain and Behavior*, *7*(4), e00660.

Roberts, R. J., Hager, L. C., &Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. *Journal of Experimental Psychology: General, 123,* 374–393.

Rogalski, Y., Peelle, J. E., & Reilly, J. (2011). Effects of perceptual and contextual enrichment on visual confrontation naming in adult aging. *Journal of Speech, Language, and Hearing Research*, *54*(5), 1349-1360.

Salthouse. T.A., & Meinz, E.J. (1995). Aging, inhibition, working memory, and speed. *Journal of Gerontology: Psychological Sciences, 50B,* P297-P306.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*(3), 403.

Salthouse, T. A., & Ferrer-Caja, E. (2003). What needs to be explained to account for age-related effects on multiple cognitive variables?. *Psychology and Aging*, *18*(1), 91.

Schlichting, L. (2005). Peabody picture vocabulary test-III-NL. *Amsterdam, the Netherlands: Hartcourt Assessment BV.*

Schubert, A. L., Hagemann, D., Voss, A., Schankin, A., & Bergmann, K. (2015). Decomposing the relationship between mental speed and mental abilities. *Intelligence*, *51*, 28-46.

Schubert, A. L., Hagemann, D., & Frischkorn, G. T. (2017). Is general intelligence little more than the speed of higher-order processing?. *Journal of Experimental Psychology: General*, *146*(10), 1498.

Shaul, S., & Nevo, E. (2015). Different speed of processing levels in childhood and their contribution to early literacy and reading abilities. *Early Childhood Research Quarterly*, *32*, 193-203.

Sheppard, L.D., Vernon, P.A. (2008). Intelligence and speed of information-processing: A review of 50 years of research. *Personality and Individual Differences, 44* (3), pp. 535-551.

Soble, J. R., Sordahl, J. A., Critchfield, E. A., Highsmith, J. M., González, D. A., Ashish, D., Marceaux, J. C., O'Rourke, J. J. F., & McCoy, K. J. M. (2016). Slow and Steady Does Not Always Win the Race: Investigating the Effect of Processing Speed across Five Naming Tests. *Archives of Clinical Neuropsychology*, *31*(8), 976-982.

Soveri, A., Lehtonen, M., Karlsson, L. C., Lukasik, K., Antfolk, J., & Laine, M. (2018). Test–retest reliability of five frequently used executive tasks in healthy adults. *Applied Neuropsychology: Adult*, *25*(2), 155-165.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*(1), 137-149.

Strijkers, K., Costa, A., & Thierry, G. (2009). Tracking lexical access in speech production: electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex, 20*(4), 912-928.

Taft, M. (2015). The nature of lexical representation in visual word recognition. In A. Pollatsek & R. Treiman (Eds.), *The Oxford Handbook on Reading* (pp. 99–113). New York: Oxford University Press.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632-1634.

Trapman, M., van Gelderen, A., van Steensel, R., van Schooten, E., & Hulstijn, J. (2014). Linguistic knowledge, fluency and meta-cognitive knowledge as components of reading comprehension in adolescent low achievers: differences between monolinguals and bilinguals. *Journal of Research in Reading, 37*(S1), S3-S21.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *The Quarterly Journal of Experimental Psychology, 67*(6), 1176-1190.

Valente, A., & Laganaro, M. (2015). Ageing effects on word production processes: an ERP topographic analysis. *Language, Cognition and Neuroscience, 30*(10), 1259-1272.

Verhaegen, C., & Poncelet, M. (2013). Changes in naming and semantic abilities with aging from 50 to 90 years. *Journal of the International Neuropsychological Society, 19*(2), 119-126.

Vernon, P. A. (Ed.). (1987). *Speed of information-processing and intelligence*. Westport, CT, US: Ablex Publishing.

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale, 4th Edition (WAIS–IV)*. San Antonio, TX: The Psychological Corporation.

Welcome, S. E., & Joanisse, M. F. (2014). Individual differences in white matter anatomy predict dissociable components of reading skill in adults. *NeuroImage*, *96,* 261-275.

Yap, M. J., Tse, C. S., & Balota, D. A. (2009). Individual differences in the joint effects of semantic priming and word frequency revealed by RT distributional analyses: The role of lexical integrity. *Journal of Memory and Language*, *61*(3), 303-325.

Appendix A – Picture naming stimuli

Target	Translation	ZipfF	Prevalence	ND	Familiarity	Visual complexity
appel	apple	4.01	1.00	2	4.9	1.7
*asperge	asparagus	2.20	0.99	1	3.8	2.2
brief	letter	4.87	1.00	2	3.7	3.5
bureau	bureau	4.83	0.99	3	3.4	3.8
cactus	cactus	3.26	0.99	0	3.5	3.9
deur	door	5.39	1.00	12	5.0	1.6
egel	hedgehog	2.76	1.00	1	3.2	3.4
eiland	island	4.71	1.00	7	2.6	3.7
emmer	bucket	3.83	1.00	6	4.0	2.1
fluit	flute	3.87	1.00	7	3.4	2.7
gieter	watering can	2.51	0.99	3	4.0	2.1
glijbaan	slide	2.70	1.00	0	3.5	3.6
helm	helmet	4.04	1.00	5	2.8	3.5
*jongen	boy	5.64	1.00	7	4.2	3.6
jurk	dress	4.75	1.00	3	4.2	2.0
koffer	suitcase	4.53	1.00	3	3.6	2.3
konijn	rabbit	4.28	1.00	2	3.9	2.8
kopje	cup	4.28	1.00	7	4.5	2.0
ladder	ladder	3.99	1.00	8	4.0	2.5
neus	nose	4.85	1.00	8	4.7	2.1
paspoort	passport	4.27	1.00	1	3.7	2.6
passer	compass	2.04	0.99	4	3.5	3.3
pinguin	penguin	3.39	0.98	0	2.8	3.3
*racket	racket	2.96	0.99	2	3.5	2.2
*rasp	grater	2.15	1.00	4	3.9	3.0
schaap	sheep	3.82	1.00	10	3.7	3.1
schoen	shoe	4.13	0.99	4	4.4	2.5
schommel	swing	3.20	0.99	2	3.9	2.4
sleutel	key	4.91	1.00	1	5.0	2.4
snavel	beak	3.24	1.00	1	2.6	2.9
tafel	table	4.92	1.00	3	4.8	1.9
trommel	drum	3.27	1.00	3	2.9	4.0
*uier	udder	1.85	1.00	16	3.0	3.7
varken	pig	4.39	1.00	6	3.6	2.8
vergiet	colander	2.94	1.00	9	4.3	2.1
vinger	finger	4.46	1.00	4	4.9	2.7
vogel	bird	4.51	1.00	2	3.1	3.8
vuist	fist	3.83	1.00	7	4.6	2.8
waaier	fan	3.01	0.99	5	2.8	3.3
wortel	carrot	3.79	1.00	2	4.7	1.8

*Note: * removed due to low name agreement. ZipfF = Zipf frequency; ND = Neighborhood density.*

Appendix B – Lexical decision stimuli

Target	Translation	ZipfF	Prevalence	ND
aardbei	strawberry	3.19	1.00	0
ambtenaar	civil servant	3.58	1.00	0
auto	car	5.66	1.00	2
bijbel	bible	4.36	1.00	1
bliksem	lightning	4.03	0.99	0
bloem	flower	4.13	1.00	6
brein	brain	4.22	1.00	9
bril	glasses	4.39	1.00	5
cello	cello	3.22	0.99	3
chocolade	chocolate	4.14	1.00	1
doorn	thorn	3.30	1.00	2
fles	bottle	4.66	1.00	2
gebouw	building	4.83	1.00	1
geraamte	skeleton	2.85	1.00	1
geschenk	gift	4.27	1.00	0
gesp	buckle	2.98	0.99	1
gitaar	guitar	4.06	1.00	0
gong	gong	2.68	0.97	10
gras	grass	4.27	1.00	12
hagel	hail	3.13	1.00	2
hark	rake	3.18	0.99	9
haver	oat	2.95	1.00	5
huig	uvula	2.04	0.99	7
jager	hunter	4.05	1.00	4
kastanje	chestnut	2.48	0.99	1
kasteel	castle	4.44	1.00	0
kegel	pin	2.63	0.99	5
kever	beetle	3.34	1.00	6
klarinet	clarinet	3.01	0.99	0
klomp	clog	2.95	1.00	4
korf	basket	2.63	1.00	2
krater	crater	3.33	1.00	4
kruis	cross	4.32	1.00	10
lamp	lamp	4.14	0.99	7
leeuw	lion	4.17	1.00	1
magneet	magnet	3.20	1.00	1
marmot	marmot	2.63	0.99	0
meeuw	gull	2.81	1.00	4
microfoon	microphone	4.01	1.00	1
pantoffel	slipper	2.36	1.00	1
papier	paper	4.49	1.00	0
parfum	perfume	4.04	0.99	1
pedaal	pedal	2.95	1.00	0

pinda	peanut	3.33	1.00	0
pistool	pistol	5.01	1.00	0
pizza	pizza	4.39	1.00	0
raket	rocket	4.17	1.00	3
rozijn	raisin	2.48	0.99	2
schemer	dusk	2.57	0.99	1
scherm	screen	4.14	1.00	3
schors	bark	3.06	1.00	3
soldaat	soldier	4.72	1.00	0
televisie	television	4.35	1.00	1
trein	train	4.86	0.99	6
veter	shoelace	3.12	0.99	7
vijg	fig	2.36	0.99	3
vlaai	flan	2.26	1.00	1
vlees	meat	4.79	1.00	4
vliegtuig	airplane	4.95	1.00	0
voetbal	football	4.11	0.99	0

Note: ZipfF = Zipf frequency; ND = Neighborhood density.

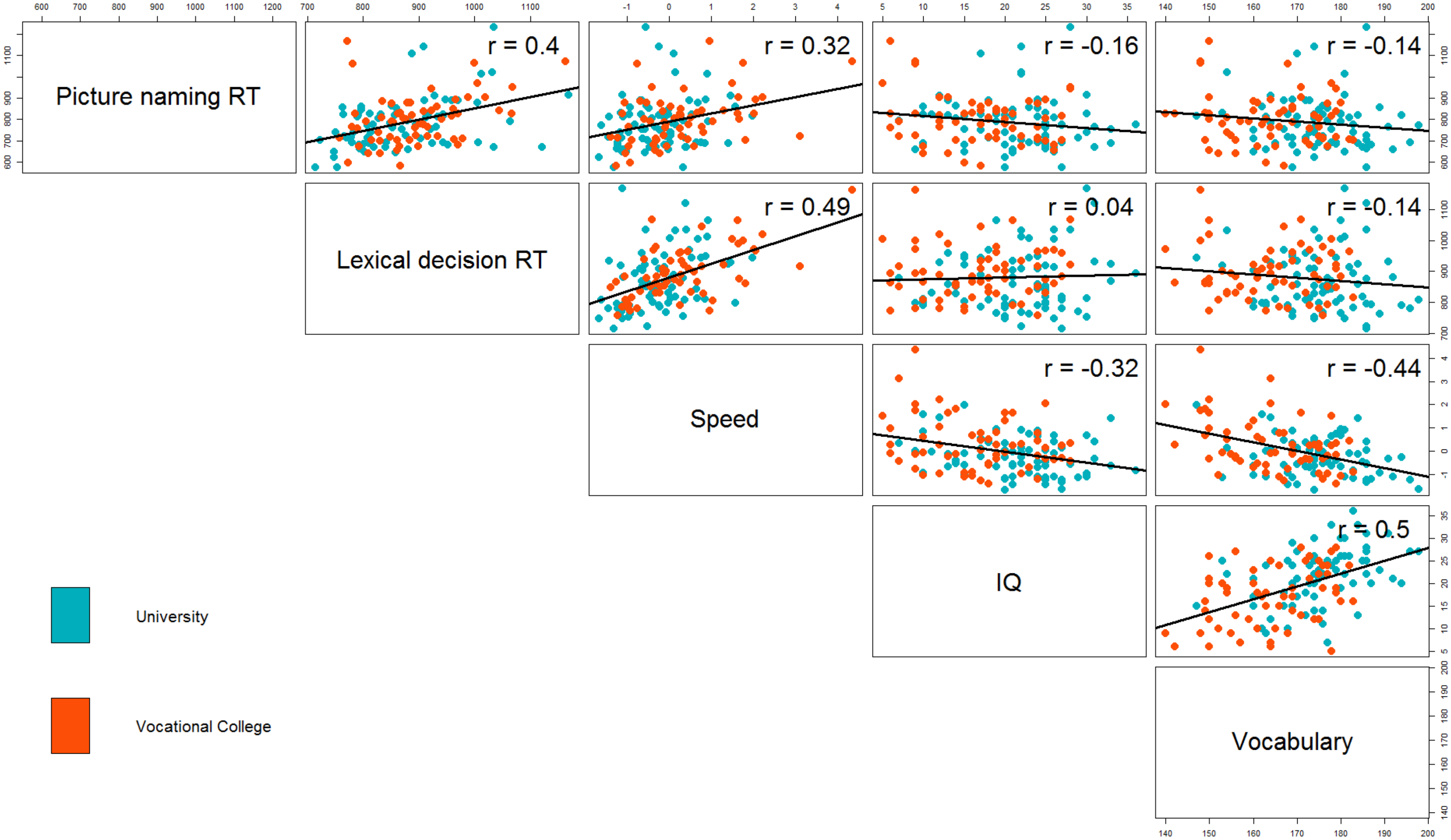

Figure 1. Correlation matrix for participants' picture naming and positive lexical decision RTs, derived processing speed factor, non-verbal intelligence and vocabulary scores