

The role of imagery-related properties in picture naming: A newly standardized set of 360 pictures for Japanese

Takehiko Nishimoto · Takashi Ueda · Kaori Miyawaki ·
Yuko Une · Masaru Takahashi

Published online: 6 January 2012
© Psychonomic Society, Inc. 2011

Abstract Picture naming was investigated primarily to determine its dependence on certain imagery-related variables, with a secondary aim of developing a new set of Japanese norms for 360 pictures. Pictures refined from the original Nishimoto, Miyawaki, Ueda, Une, and Takahashi (Behavior Research Methods 37:398–416, 2005) set were used. Naming behaviors were measured using four imagery-related measures (imageability, vividness, image agreement, and image variability) and four conventional measures (naming time, name agreement, familiarity, and age of acquisition), as well as a number of other measures (17 total). A simultaneous multiple regression analysis performed on naming times showed that the most reliable predictor was *H*, a measure of name diversity; two image-related measures (image agreement and vividness) and age of acquisition also contributed substantially to the prediction of naming times. The accuracy of picture naming (measured as name agreement) was predicted by vividness, age of acquisition, familiarity, and image agreement. This suggests that certain

processes involving mental imagery play a role in picture naming. The full set of norms and pictures may be downloaded from <http://www.psychonomic.org/archive/> or along with the article from <http://www.springerlink.com>.

Keywords Picture naming · Imageability · Vividness · Image agreement · Image variability · Name agreement · Familiarity · Age of acquisition

Picture naming is one of the most commonly used paradigms in the field of psycholinguistics; nevertheless, the activity underlying picture naming is complex, in that it is affected by a number of different factors (e.g., visual complexity, familiarity, name agreement, word frequency, and age of acquisition) in the various levels of picture processing, from object recognition to articulation. Therefore, it is essential to establish norms for pictorial stimuli related to these factors in order to carry out the study of picture naming. The pioneering study of Snodgrass and Vanderwart (1980) examined picture naming using a standardized set of 260 pictures in American English. Over the following years, similar reports have addressed picture naming in French (Alario & Ferrand, 1999; Bonin, Peereman, Malardier, Méot, & Chalard, 2003), Spanish (Sanfeliu & Fernandez, 1996), Argentinean Spanish (Manoiloff, Artstein, Canavoso, Fernández, & Segui, 2010), Icelandic (Pind, Jónsdóttir, Gissurardóttir, & Jónsson, 2000; Pind & Tryggvadóttir, 2002), Italian (Dell'Acqua, Lotto, & Job, 2000; Lotto, Surian, & Job, 2010), Greek (Dimitropoulou, Duñabeitia, Blitsas, & Carreiras, 2009), British English (Johnston, Dent, Humphreys, & Barry, 2010), Dutch (Severens, Van Lommel, Ratinckx, & Hartsuiker, 2005), Japanese (Matsukawa, 1983; Nishimoto & Hayashi, 1996; Nishimoto, Miyawaki, Ueda, Une, & Takahashi, 2005; Nishimoto & Yasuda, 1982), and Chinese (Weekes, Shu, Hao, Liu, & Tan, 2007). To investigate the factors that determine cross-language universals and cross-language

The authors gratefully acknowledge Editor Gregory Francis for his encouraging advice, and we thank anonymous reviewers for their constructive and helpful suggestions for the manuscript.

Electronic supplementary material The online version of this article (doi:10.3758/s13428-011-0176-7) contains supplementary material, which is available to authorized users.

T. Nishimoto · T. Ueda · K. Miyawaki · Y. Une
Waseda University,
Tokyo, Japan

M. Takahashi
Foundation Study Center, Saitama Institute of Technology,
Saitama, Japan

T. Nishimoto (✉)
Department of Psychology, Faculty of Letters, Arts and Sciences,
Waseda University,
1-24-1 Toyama, Shinjuku-ku,
162-8644, Tokyo, Japan
e-mail: nishi@waseda.jp

disparities, Bates et al. (2003) collected object-naming norms for 520 line drawings of common objects in seven languages: American English, Spanish, Italian, German, Bulgarian, Hungarian, and Mandarin Chinese. In addition to sets of line drawings, normative sets of photographs are now available (Viggiano, Vannucci, & Righi, 2004, with 174 photos; Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010, with 480 photos). In the present study, we provide a refined set of Japanese norms for pictures and investigate the processes involved in picture naming by examining the individual contributions of a number of factors to naming times and accuracy, focusing specifically on imagery-related factors.

Numerous variables have been proposed to characterize picture-naming time and accuracy. Table 1 gives a list of commonly used variables (all of them adopted for the present study), along with brief definitions.

In referring to models of picture naming (e.g., Humphreys, Riddoch, & Quinlan, 1988; Levelt, Roelofs, & Meyer, 1999), Alario et al. (2004) discussed empirical findings related to the speed of picture naming. They suggested that several processes may determine response speed that depends on a number of different factors. Among the determinants of theoretical interest were visual factors (such as visual complexity and image agreement), semantic factors (such as conceptual familiarity and imageability), lexical factors (such as

frequency and age of acquisition), and lexicalization and articulation. With respect to name agreement (or verbal codability), which refers to the degree to which all participants agree on the name for a given picture, its impact on the time to identify picture names has been ascribed to either (1) difficulties in assessing stored structural representations when objects are difficult to identify (e.g., a spider vs. an ant) or (2) difficulties in selection occurring after semantic access, due to different, but potentially correct, names associated with the same object (e.g., *couch* vs. *sofa*; Bonin, Chalard, Méot, & Fayol, 2002; Vitkovitch & Tyrrell, 1995). A number of previous studies have attempted to identify factors that contribute to picture-naming time (e.g., Alario et al., 2004; Barry, Morrison, & Ellis, 1997; Bonin, Barry, Méot, & Chalard, 2004; Bonin et al., 2002; Cuetos & Alija, 2003; Ellis & Morrison, 1998; Lotto, Surian, & Job, 2010; Paivio et al., 1989; Snodgrass & Yuditsky, 1996). Table 1 shows the various factors that might be expected to have some effect on naming times. With the exception of a very few studies, much of this research has not systematically focused on imagery-related factors (e.g., imageability, image variability, and vividness) as possible determinants of picture-naming responses (i.e., accuracy and naming times). Bonin et al. (2002) found that one of the major determinants of both written and spoken picture-naming times was image variability; Bonin et al.

Table 1 List of the measures and their abbreviations

Measure	Abbreviation	Definition of Measure	Task
Response time	RT _{st}	Picture-naming times calculated with respect to the strict criteria	Timed picture naming
	RT _{lib}	Picture-naming times calculated with respect to the liberal criterion	
Naming agreement	NA _{st}	Degree to which all participants agreed on the name for a given picture with respect to the strict criteria	
	NA _{lib}	Degree to which all participants agreed on the name for a given picture with respect to the liberal criterion	
Naming diversity	<i>H</i>	Diversity or range of the collection of obtained names	
Name agreement	uNA _{st}	Degree to which all participants agreed on the name for a given picture with respect to the strict criteria	Untimed picture naming
	uNA _{lib}	Degree to which all participants agreed on the name for a given picture with respect to the liberal criterion	
Naming diversity	u <i>H</i>	Diversity or range of the collection of obtained names	
Image agreement	IA	Degree to which each picture resembled participants' mental images of the object	IA rating
Image variability	IV	Degree to which a name evoked different images for a particular object	IV rating
Imageability	IMG	Ease and speed with which a target word evoked a corresponding mental image	IMG rating
Vividness	VIV	Perceived clarity and amount of detail of a mental image generated from a word	VIV rating
Familiarity	FAM	Familiarity of a concept	FAM rating
Age of acquisition	AoA	Age at which names were acquired	AoA rating
Complexity	CMP	Visual complexity, approximated by the size of the JPEG-formatted file of a picture	—
Log frequency	LogFREQ	Printed frequency with which a name had appeared in the <i>Asahi Shimbun</i> (a major newspaper in Japan) across 14 years	—
Mora	MORA	Number of Japanese speech units, which approximately correspond to syllables in English	—

(2004) then examined the influences of both imageability and concreteness, as well as of conceptual familiarity. However, no clear picture of the influence of imagery-related factors has emerged, because the relationship between the various factors that contribute to naming times and the processing levels at which they putatively function is not straightforward. Indeed, this topic awaits further investigation.

Originally, Snodgrass and Vanderwart (1980) defined two imagery measures in picture naming, namely image agreement and image variability. In terms of the current models of picture naming, mentioned above, image agreement has been proposed to exert an influence on naming behavior at the processing level of visual recognition, whereas image variability is considered to impact the semantic level of coding. The former is based on findings in tasks in which participants are asked to judge “how closely each picture resembles their mental image of the object,” and the latter on tasks in which participants are asked “whether the name evokes few or many different images for that particular object.” Bonin et al. (2002) found that these two measures were statistically significant as major determinants of naming time. This finding suggests that those picture objects that elicit both high image agreement and high image variability may possess semantic richness or strength.

Imageability is a semantic property that is defined as the ease and speed with which a target word evokes a corresponding mental image (Clark & Paivio, 2004; Paivio, Yuille, & Madigan, 1968). It has been suggested that words with high imageability lead to better memory performance than do those with low imageability in both recall and recognition tasks, and that imageability strongly determines responding in word association tasks (e.g., de Groot, 1989). Words with high imageability are also translated more quickly and accurately than are abstract words (e.g., van Hell & de Groot, 1998). Furthermore, several psycholinguistic tests, such as PALPA (Psycholinguistic Assessments of Language Processing in Aphasia; Kay, Lesser, & Coltheart, 1992) include imageability as a major variable in their test batteries. Moreover, several studies using neuroimaging have suggested effects due to imageability in cognitive processing (e.g., Frost et al., 2005; Jefferies, Patterson, Jones, & Ralph, 2009). It is an important variable for cognitive studies in fields that deal with memory, language, or higher-order functional deficits of the brain.

A series of comprehensive studies examining the imageability of words began with Paivio et al. (1968). Their participants were asked to judge the ease or difficulty with which they could generate a mental image for each of 925 words. This study was followed by various studies of the same kind in American English (Coltheart, 1981), Dutch (van Loon-Vervoorn, 1985), British English (Morrison, Chappell, & Ellis, 1997), Spanish (Izura, Hernández-Muñoz, & Ellis, 2005), Canadian French (Desrochers & Bergeron, 2000), Japanese (Sakuma, Ijuin, Fushimi,

Tatsumi, & Tanaka, 2005), and French (Desrochers & Thompson, 2009). The measure of imageability for words may be a significant variable in the prediction of word-naming times (Strain, Patterson, & Seidenberg, 1995). For picture naming, the influence of imageability remains undetermined. Lloyd-Jones and Nettlemill (2007) treated imageability as an important variable largely because it was shown to significantly affect picture-naming times. On the other hand, Morrison, Ellis, and Quinlan (1992) failed to find a significant effect of imageability on picture naming. Ellis and Morrison (1998) also found no significant effects using standard regression procedures, although a reanalysis using another procedure did reveal significant effects. Barry et al. (1997) found that the influence of imageability on naming times would vary according to the other predictors involved in multiple regression analyses. Bonin et al. (2004) also examined the influence of imageability, as well as of other variables, and reported a reliable effect of imageability. Cuetos and Alija (2003) could not find a significant effect of imageability in determining predictors of naming times for action pictures.

Vividness is another concept that includes an imagery-related property. It is based on findings in tasks in which participants are asked to judge how clear the perceived mental image generated from the word is. The measure represents the perceived clarity and amount of detail of a mental image; that is, it captures the image strength or the extent to which visual information is specific and rich (D’Angiulli, 2003–2004). Although vividness is highly correlated to imageability (e.g., $r = .83$ for the 925 words in Paivio et al., 1968), D’Angiulli (2003–2004) claimed that the effect of vividness is distinct from that of imageability. He collected word-naming times and ratings of vividness for 150 nouns. The imageability, concreteness, and meaningfulness norms for these words were borrowed from Paivio et al. (1968). He demonstrated that word-naming time was strongly related to vividness when imageability was controlled. Conversely, naming time was strongly related to imageability for words that had approximately the same vividness. D’Angiulli argued that imageability may be a correlate of the activation of frontal regions, reflecting the degree of voluntary control and willful effort that is being exerted while generating mental visual images. In contrast, vividness may be a correlate of the activation of association visual areas, reflecting processes involved in the inspection or evaluation of the retrieved sensory information contained in the images. Furthermore, with regard to working memory, he suggested that imageability may reflect processes of the central executive, and that the main processing functions that correspond to vividness may depend on the visuospatial sketchpad. Moreover, D’Angiulli and Reeves (2007) examined how the relationship between vividness and word-naming times might reflect the action of the two visual imagery pathways hypothesized by Kosslyn (1994): the ventral pathway, which

processes object properties, and the dorsal pathway, which processes locative properties of mental images. In the present approach, we distinguish vividness, as defined by D'Angiulli (2003–2004), from imageability, as defined by Paivio et al. (1968), and adopt both as imagery-related measures. We have also included two conventional measures—namely, image agreement and image variability. Therefore, in total we obtained four imagery-related measures.

The pictures for the present approach were taken from the Nishimoto et al. (2005) image set. However, we recognized the possibility of generational differences across extended periods of time for some of these pictures (e.g., *tape recorder*, *microwave*, and *traffic light*); accordingly, some pictures required improvement in regard to name agreement. Hence, we redrew 82 pictures and added 1 picture to the prior set, resulting in 360 pictures in total for the present set.

Using this set, the naming times for each of the 360 pictures were measured. Additionally, seven predictors of picture-naming time were collected: name agreement, familiarity, age of acquisition, imageability, vividness, image agreement, and image variability. As indicated above, imageability, vividness, image agreement, and image variability represent new measures in the present approach. In the picture-naming task, participants were instructed to identify each picture as briefly as possible within a given period. In addition to this timed naming, we introduced an untimed naming task in which participants were allowed to respond without a time limit. Time pressure might influence the processes involved in picture naming, which then would be reflected in name agreement. Snodgrass and Yuditsky (1996) pointed out that the pattern of naming errors varies according to whether responses are untimed or timed, while Székely et al. (2003) found that the basic measures (e.g., name agreement and *H* statistics) obtained in timed picture naming are highly correlated with those in untimed picture naming. In the present study, we introduced both timed and untimed naming tasks in order to further evaluate the influence of time pressure for responses.

Because a considerable number of additional measures (described in the [Method](#) section) were adopted, the total number of measures was 17. These measures included four imagery-related measures (imageability, vividness, image agreement, and image variability), in addition to four conventional measures (naming times, name agreement, familiarity, and age of acquisition) and other measures of lexical properties.

In brief, the present study had two aims. One was to investigate whether picture naming might be predicted by one or several variables, with a special interest in the imagery-related variables. The predictions were that at least naming agreement and age of acquisition would be strong determinants of naming performance, because both of these variables have been found to be highly important predictors

in previous picture-naming studies. In addition, we focused on the effects of the image-related variables on naming speed. The second aim was to provide a refined set of Japanese norms for the 360 pictures. [Appendix A](#) of the supplemental materials lists the norms that were developed in this study. The pictures with lower name agreement in the Nishimoto et al. (2005) set were redrawn to improve name agreement.

Method

Participants

A total of 1,217 native speakers of Japanese voluntarily participated in the tasks. Almost all of the participants were undergraduates. No participants took part in more than one task in the experimental session. The numbers of participants differed for tasks designed to assess the eight different measures of interest; these numbers were as follows: 121 in the timed naming task, 458 in the untimed naming task, 112 in the age-of-acquisition rating, 119 in the image variability rating, 94 in the familiarity rating, 117 in the imageability rating, 89 in the image agreement rating, and 107 in the vividness rating. Of the 107 participants in the vividness rating, 1 participant was omitted for not submitting any responses.

Stimuli

The stimuli for the naming tasks and the image agreement rating consisted of 360 line drawings. Of these pictures, 216 were identical to the Snodgrass and Vanderwart (1980) pictures. Nishimoto et al. (2005) redrew 44 pictures from the Snodgrass and Vanderwart set and added 99 new pictures. From these 143 pictures, we redrew 82 pictures that had relatively low name agreement scores in the Nishimoto et al. (2005) set and added 1 new item. [Appendix B](#) of the supplemental materials shows all of the pictures that were either redrawn or added.

The names for each picture were determined by the timed naming task. The most frequent name given in the task was defined as a dominant name for that picture. These dominant names were used in the ratings of familiarity, age of acquisition, imageability, image agreement, image variability, and vividness.

The 360 pictures were numbered from 1 to 360 according to the order of the first syllable of their names in the Japanese kana order (syllabary). To reduce the task durations for participants, we divided the total set of pictures into three subsets of 120 pictures each on the basis of the picture numbers, in accordance with Snodgrass and Yuditsky's (1996) procedure: If the remainder was 1 when the picture

number was divided by 3, the picture was assigned to Set 1; if the remainder was 2, the picture was assigned to Set 2; otherwise, the picture was assigned to Set 3. All of the participants, except those who took part in the untimed naming task, rated (or named) one of the sets. All pictures except those taken unmodified from Snodgrass and Vanderwart (1980) are provided in Appendix B.

Apparatus

In the timed naming task, a personal computer was used to control presentation of the stimuli; responses were collected with a voice key connected to the PC. In the image agreement rating, a projector connected to a PC was used to present the stimuli, and booklets were used to fill in the rating scores.

In the other rating tasks, custom-printed booklets were used. Each booklet consisted of 120 items in a set and a corresponding rating scale. The presentation order of the items was randomized separately for each copy of the booklet used in every rating task. In each rating task, booklets were randomly assigned to the participants.

In the untimed naming task, booklets were also used. In this task, all 360 pictures were randomly divided into 18 subsets of 20 pictures each, regardless of their position in the three sets mentioned before. The 20 pictures were randomly ordered separately for each printed copy of these booklets.¹

Procedure

Details for each of the eight tasks in which the 17 measures were derived are outlined below.

Timed naming (RT_{st} , RT_{lib} , NA_{st} , NA_{lib} , H)

The timed naming procedure was identical to that in Nishimoto et al. (2005). The participants were asked to articulate the name of the randomly presented picture as quickly and accurately as possible. Each participant named 120 pictures in a particular set, and the participants were randomly assigned to one of the sets. Responses with response times (RTs) under 50 ms or exceeding 10 s were regarded as ERROR responses, in line with Snodgrass and Yuditsky (1996); these resulted mainly from insufficient functioning of the voice key. The RTs of participants were recorded via voice key, and the verbal responses were simultaneously written down by the experimenter.

¹ The participants were asked to indicate the category to which each stimulus belonged, as well as the typical colors and features of the presented pictures and names. The category, typical-color, and feature data were collected for another study.

Defining dominant names After naming responses were obtained, the most frequent response to a given picture was defined as the dominant name for that picture. The naming correctness of the pictures was determined on the basis of these lists of dominant names. Although there were timed and untimed dominant names for a particular picture, the former were used as the stimulus words in the rating tasks, which was in accordance with Nishimoto et al. (2005).

Two criteria for naming accuracy Naming accuracy was defined using two different criteria—namely, a *strict* and a *liberal* criterion. We judged a named response as *strictly* correct if it conformed to one of the following criteria:

- (1) It was the same as a dominant name.
- (2) It was part of a dominant name (e.g., *shirt* for *dress shirt* or *finger* for *index finger*).
- (3) It was an abbreviated word starting with the same phoneme as a dominant name (e.g., *heri* for *herikoputâ* “helicopter”).
- (4) It was the first word of a dominant name including two or more words (e.g., *ashi* for *ashi-no-yubi* “toe”).
- (5) It was an idiomatic name subsuming the dominant name (e.g., *happa* for *ha* “leaf,” or *chouchou* for *chou* “butterfly”).
- (6) It was the more authentic name (e.g., *seiyo-nashi* for *nashi* “pear”).

Regardless of the above criteria, a named response was viewed as *liberally* correct when at least 2 participants named the same word for a picture, even if this was a wrong answer.

RTs for each of the 360 pictures were averaged over participants’ correct responses using the same two criteria for determining accuracy—that is, the strict criteria (RT_{st}), and the liberal criterion (RT_{lib}). RT_{st} were the mean RTs for only the strictly correct responses to an item, whereas RT_{lib} were the means when liberally correct responses were included as well.

In addition to the RT measures, we calculated two measures of name accuracy— NA_{st} for naming agreement under the strict criteria and NA_{lib} for naming agreement under the liberal criterion—and H for the diversity or range of the collection of obtained names as a whole. NA_{st} and NA_{lib} represented the percentages of participants giving the correct answer under the strict and liberal criteria, respectively. Naming diversity (H)

was defined as the entropy function $H = \sum_{i=1}^k P_i \log_2(1/P_i)$ (Snodgrass & Vanderwart, 1980), calculated from the proportion (P) of each name among the k names given for each picture. Because H does not take into account the “correctness” of responses, neither the strict nor the liberal criteria was applied; hence, H represents simple naming diversity rather than correctness.

Untimed naming (uNA_{st} , uNA_{lib} , uH)

The participants were asked to report the names of the pictures in their booklets.

Identifying dominant names for untimed naming Dominant names were defined as for timed naming. Even if the written responses from the participants assumed several different forms of notation (e.g., Japanese kanji, katakana, and hiragana characters, as well as some misspelled names), we did not separately evaluate different forms of the same response; they were treated as a single named responses.

Criteria for naming accuracy and measures for name agreement We adopted the same criteria (strict, liberal) as in the timed naming task. The measures for name agreement in the untimed tasks were also the same as in timed naming. These are uNA_{st} , uNA_{lib} , and uH , where the prefix “u” designates “untimed.”

Appendix A of the supplemental materials provides the dominant names for each item obtained in both timed and untimed naming tasks, as well as the measures and statistics for naming time and agreement. However, the names for the concepts used in the other rating tasks were the timed ones.

Familiarity rating (FAM)

In this task, participants rated the familiarity of the concept represented by each word. They were instructed to rate, on a 7-point scale, how familiar each concept described by the name shown in the booklet was in each participant’s experience; 1 indicated *extremely unfamiliar* and 7 indicated *extremely familiar*. If participants did not know an object, they were asked to respond “don’t know the object” (DKO). They were also instructed not to rate the presented word, but the concept itself.

Age of acquisition rating (AoA)

Following Carroll and White (1973a, 1973b), participants were instructed to estimate their age (in years) when they had learned the concept represented by a word, using a 9-point scale (where the 1 to 9 ratings represented 2 years, 3 years, 4 years, 5 years, 6 years, 7–8 years, 9–10 years, 11–12 years, and 13 years or older, respectively). It was emphasized that their rating should apply to the acquisition of the concept, not the word.

Imageability rating (IMG)

The participants rated the imageability of the dominant names (Clark & Paivio, 2004; Paivio et al., 1968). The

participants were instructed to judge how quickly and easily each word generated a mental image on a 7-point scale, from 1 (*extremely difficult*) to 7 (*extremely easy*). If they did not know the object, they were asked to respond “don’t know the object.” Booklets presenting 120 of the names from the set and the imageability scales were given to the participants.

Vividness rating (VIV)

The participants rated the vividness for the concept presented by the word. Booklets presenting 120 of the dominant names from the set and the vividness scales were given to the participants. The participants were instructed to judge the clarity of the perceived mental image generated from the word on a 7-point scale from 1 (*extremely unclear; hard to have an image to the word*) to 7 (*extremely clear: the same as seeing the actual object*). If the participants did not know the word, they were asked to respond “don’t know the name” (DKN)

Image agreement rating (IA)

Participants generated a mental image from the word presented on the screen and rated how closely it resembled the picture. The rating procedure was an approximation of that used in Snodgrass and Vanderwart (1980). Participants rated a set of 120 name–picture pairs. In each trial, a dominant name was projected on the screen for 5 s. In this period, the participants generated a mental image of the object shown by the word. Afterward, the picture for that word was displayed for 5 s. The participants were asked to rate the extent of the resemblance of the projected image to the mental images they had formed from the names on a 7-point scale, in which 1 indicated *low agreement* and 7 *high agreement* (i. e., resembled very closely). The name–picture pairs were displayed on the screen in random order. If the participants did not form images from a word, they were asked to respond “can’t imagine,” and if they imagined a different object from the word (e.g., the participant formed an image of a homonym of the word), they were asked to respond “different object.”

Image variability rating (IV)

The participants rated how many different mental images could be evoked by the presented words using a 7-point scale, in which 1 indicated *extremely few* and 7 *very many*. This rating roughly followed the procedure in Snodgrass and Vanderwart (1980). Participants were given a booklet containing 120 dominant names in a random order and were instructed to estimate how many different mental images they generated in response to each word in the booklet. If

the participants did not know a word, they were asked to respond “don’t know the name,” and if no images came to mind, they were to respond “can’t form images.”

Other indices (CMP, LogFREQ, MORA)

We also included measures concerning the physical or lexical properties of the items in our data analysis. The complexity of the pictures (CMP) was defined as the JPEG-formatted file size of each picture file (in kilobytes; see Székely & Bates, 2000; Székely et al., 2003). Word frequency (LogFREQ) was taken from the norms of Amano and Kondo (2000), one of the most exhaustive corpora for Japanese naming words, consisting of around 13.9 million sentences (equal to around 1.2 GB of text data). Because some of the names were missing from these norms, analyses including LogFREQ were carried out with only 211 concepts. The number of morae (MORA), a Japanese speech unit that approximately corresponds to a syllable in English, was counted for each dominant name.

Results and discussion

Summary statistics for the entire set of measurements are shown in Table 2. It is notable that the imageability (IMG) of each picture was highly rated ($M = 6.17$ on a 7-point scale) with low variance ($SD = 0.19$; see Table 2). This high rating was due to the fact that the concepts used in this normative set were concrete and common in everyday life, whereas the norms of Paivio et al. (1968) included abstract concepts.

Naming failure in timed naming

We collected 14,520 responses over all participants. We identified four basic types of naming failures:

1. ERROR, which were malfunctions of the instruments and outliers in the response data, where the latter were defined as RTs of less than 50 ms or more than 10 s; errors comprised 0.68% of the responses (99 responses).
2. DKO (“don’t know object”), where the participants did not know what the picture depicted; 1.32% of the responses (191 responses) fell into this category.
3. DKN (“don’t know name”), where the participants recognized the object depicted but did not know its name; these made up 2.42% of the responses (351 responses).
4. TOT (“tip of the tongue”), where the participants knew the object but the name was on “the tip of the tongue” and could not be recalled within the allotted time; these made up 1.85% of the responses (269 responses).

Table 2 Summary of descriptive statistics

	Mean	SD	Min	Max
RT _{st} (ms)	1,175.63	338.37	685.08	2,663.00
RT _{lib} (ms)	1,192.15	346.27	685.08	2,741.75
NA _{st} (%)	81.53	21.17	7.32	100.00
NA _{lib} (%)	91.34	13.42	20.00	100.00
<i>H</i>	0.70	0.69	0.00	2.93
uNA _{st} (%)	84.66	18.68	23.53	100.00
uNA _{lib} (%)	95.70	6.87	45.45	100.00
u <i>H</i>	0.62	0.69	0.00	3.35
IA	4.92	1.02	1.96	6.80
IV	3.33	1.11	1.30	6.67
IMG	6.17	0.19	5.56	6.73
VIV	5.96	0.66	2.67	6.83
FAM	4.85	1.04	2.13	6.66
AoA	4.18	0.88	2.17	6.71
CMP	53.45	26.21	9.47	183.12
LogFREQ	6.64	1.57	1.95	11.02
MORA	3.29	1.27	1.00	8.00

RT_{st}, response time with respect to the strict criteria; RT_{lib}, response time with respect to the liberal criterion; NA_{st}, naming agreement with respect to the strict criteria; NA_{lib}, naming agreement with respect to the liberal criterion; *H*, naming diversity; uNA_{st}, naming agreement with respect to the strict criteria for untimed naming; uNA_{lib}, naming agreement with respect to the liberal criterion for untimed naming; u*H*, naming diversity for untimed naming; IA, image agreement; IV, image variability; IMG, imageability; VIV, vividness; FAM, familiarity; AoA, age of acquisition; CMP, visual complexity; LogFreq, log frequency of a given name; MORA, number of morae

Consequently, we used 13,610 responses, excluding the responses noted above.

The improvement of pictures

We had redrawn some of the pictures for this standardization, so we examined NA_{st} rating improvement for the redrawn pictures from Nishimoto et al. (2005)’s norms. The name agreement of the redrawn pictures in this version ($M = 75.95\%$) significantly improved from the corresponding pictures in the previous set [$M = 67.59\%$; $t(81) = 4.123$, $p < .001$].

Name agreement of timed naming and untimed naming

As described above, we collected not only timed naming, but also untimed naming. We compared name agreement in the timed and untimed naming tasks. All of the measures of name agreement and diversity for untimed naming (uNA_{st}, uNA_{lib}, and u*H*) were higher than those for timed naming (NA_{st}, NA_{lib}, and *H*; $t(359) = 4.613$, $p < .001$; $t(359) = 7.968$, $p < .001$; and $t(359) = 3.037$, $p < .01$, respectively). The dominant names of

Table 3 Correlations among measures

	RT _{st}	RT _{lib}	NA _{st}	NA _{lib}	H	uNA _{st}	uNA _{lib}	uH	IA	IV	IMG	VIV	FAM	AoA	CMP	LogFreq	MORA
RT _{st}	1.00																
RT _{lib}	.96**	1.00															
NA _{st}	-.76**	-.79**	1.00														
NA _{lib}	-.79**	-.82**	.77**	1.00													
H	.67**	.71**	-.88**	-.60**	1.00												
uNA _{st}	-.58**	-.61**	.80**	.55**	-.76**	1.00											
uNA _{lib}	-.55**	-.57**	.56**	.65**	-.50**	.55**	1.00										
uH	.64**	.68**	-.80**	-.63**	.78**	-.96**	-.70**	1.00									
IA	-.31**	-.33**	.31**	.28**	-.29**	.37**	.31**	-.40**	1.00								
IV	-.31**	-.32**	.27**	.31**	-.14*	.12*	.11*	-.11*	-.22**	1.00							
IMG	.00	-.04	.00	.02	.03	.01	-.03	-.01	.02	-.02	1.00						
VIV	-.54**	-.55**	.56**	.55**	-.44**	.48**	.44**	-.50**	.11*	.48**	.04	1.00					
FAM	-.38**	-.39**	.33**	.35**	-.20**	.24**	.22**	-.24**	-.05	.62**	.07	.71**	1.00				
AoA	.51**	.52**	-.47**	-.46**	.40**	-.36**	-.33**	.39**	.04	-.40**	-.07	-.53**	-.53**	1.00			
CMP	.05	.07	-.12*	-.13*	.08	-.06	-.09	.07	-.06	-.17**	-.05	-.27**	-.28**	.16**	1.00		
LogFreq	-.44**	-.44**	.38**	.42**	-.27**	.25**	.18**	-.23**	-.06	.58**	-.11	.41**	.52**	-.45**	-.05	1.00	
MORA	.20**	.22**	-.15**	-.21**	.12*	-.09	-.17**	.12*	.08	-.37**	.01	-.25**	-.20**	.40**	.04	-.43**	1.00

* $p < .05$. ** $p < .01$. RT_{st}, response time with respect to the strict criteria; RT_{lib}, response time with respect to the liberal criterion; NA_{st}, naming agreement with respect to the strict criteria; NA_{lib}, naming agreement with respect to the liberal criterion; H, naming diversity; uNA_{st}, naming agreement with respect to the strict criteria for untimed naming; uNA_{lib}, naming agreement with respect to the liberal criterion for untimed naming; uH, naming diversity for untimed naming; IA, image agreement; IV, image variability; IMG, imageability; VIV, vividness; FAM, familiarity; AoA, age of acquisition; CMP, visual complexity; LogFreq, log frequency of a given name; MORA, number of morae

34 pictures differed between the timed and untimed naming (e.g., *fur seal* and *earless seal*). Some picture names were more specific (e.g., *pen* vs. *ballpoint pen*) or more informal (e.g., common Japanese pronunciations such as “Asuparagasu” and “Asupara” for *asparagus*) in untimed naming. Some pictures were considered to reflect different concepts (e.g., *girl* vs. *doll*). In addition, 21 pictures exceeded 2 *SDs* from the mean difference between NA_{st} and uNA_{st} (e.g., *rolling pin*, *hat*, and *skate*).

Correlations among measures

Table 3 shows a correlation matrix including all of the measures. Almost all correlations among the variables were statistically significant. As expected, measures of name agreement and diversity (i.e., NA_{st} , NA_{lib} , H , uNA_{st} , uNA_{lib} , uH) were highly correlated with each other. The correlations of IMG (imageability) with the rest of the variables were not significant. The latter outcome was due to a ceiling effect; that is, participants consistently rated many concepts with the highest rating (7) on the IMG scale. On the other hand, VIV (vividness) measures were highly correlated with measures of name agreement, AoA, and FAM. Most of the correlations for CMP (complexity) were not significant. The correlations for MORA were relatively small.

Multiple regression analyses

A simultaneous multiple regression analysis was carried out on RT_{st} , using name diversity (H), IA, IV, IMG, VIV, FAM, AoA, CMP, LogFREQ, and MORA as predictor variables. Table 4 shows the results. The most reliable source of variance was H , followed by IA, AoA, and VIV. Two variables, IV and LogFREQ, were marginally significant, whereas four variables, IMG, FAM, CMP and MORA, were not significant.

These results indicate that the most reliable predictor was H (name diversity); this result is consistent with findings from our previous study in Japanese (Nishimoto et al., 2005) and with other preceding studies (e.g., Alario et al., 2004; Barry et al., 1997; Bonin et al., 2003; Johnston et al., 2010; Snodgrass & Yuditsky, 1996; Weekes et al., 2007). Second, IA (image agreement) also highly contributed to predict naming times (for similar results, see Bonin et al., 2002; Bonin et al., 2003). Since IA reflected similarity between the mental images evoked by concept names and the pictures representing those concepts (Snodgrass & Vanderwart, 1980), pictures with high IA might resemble canonical mental images that participants hold regarding the concepts depicted in those pictures (Alario et al., 2004; Cuetos, Ellis, & Alvarez, 1999). Therefore, it is reasonable that concepts of pictures with high IA were retrieved and named quickly. Third, AoA also reliably contributed to predict RTs. This

result, too, is consistent with the previous studies, which found AoA to be a robust predictor of naming time (e.g., Alario et al., 2004; Bonin et al., 2002; Bonin et al., 2003; Cuetos & Alija, 2003; Cuetos et al., 1999; Johnston et al., 2010; Snodgrass & Yuditsky, 1996). Fourth, FAM (familiarity) was not a reliable predictor of naming times. Some previous studies have also reported that familiarity did not contribute to prediction of naming times (e.g., Barry et al., 1997; Dell’Acqua et al., 2000; Pérez, 2007), although others have found significant effects of familiarity (e.g., Cuetos et al., 1999; Johnston et al., 2010; Snodgrass & Yuditsky, 1996). Related to these results, Bonin et al. (2004) reported a trade-off between AoA and familiarity. That is, when rated AoA and familiarity were included as predictor variables, AoA was not significant, whereas familiarity was significant. By contrast, when objective AoA and familiarity were included, AoA was significant and familiarity was not. When familiarity was included without AoA, familiarity was significant. Similarly, our results also indicated that AoA was significant but familiarity was not. These results suggest that to some extent AoA and familiarity reflect a common process. Fifth, VIV (vividness) was a significant predictor of naming times, indicating that some of the measures involving mental images (VIV and IA) contribute substantially to predictions of naming times. Finally, as in previous studies, IMG (imageability) did not make a significant contribution (e.g., Barry et al., 1997; Morrison et al., 1992; but see Bonin et al., 2004). This result might be interpreted to mean that imageability, which is assumed to

Table 4 Results of multiple regression analyses on RT_{st} and NA_{st}

	RT_{st}		NA_{st}	
	β	t	β	t
H	.42	8.24**	–	–
IA	–.28	–6.21**	.25	4.75**
IV	–.11	–1.91 ⁺	.12	1.63
IMG	.03	0.66	–.05	–1.00
VIV	–.15	–2.26*	.56	7.66**
FAM	.01	0.20	–.29	–3.44**
AoA	.19	3.46**	–.29	–4.70**
CMP	–.06	–1.49	–.01	–0.14
LogFREQ	–.11	–1.95 ⁺	.18	2.82**
MORA	.07	1.46	.16	2.86**
Adjusted R^2	.67		.52	
F value	42.75**		26.43**	

* $p < .05$. ** $p < .01$. ⁺ $p < .10$. RT_{st} , response time with respect to the strict criteria; NA_{st} , naming agreement with respect to the strict criterion; H , naming diversity; IA, image agreement; IV, image variability; IMG, imageability; VIV, vividness; FAM, familiarity; AoA, age of acquisition; CMP, visual complexity; LogFreq, log frequency of a given name; MORA, number of morae

be a semantic variable, actually captured only part of the semantic properties. However, this might also be a result of the distributional bias of IMG values, discussed above. Therefore, further research will be needed to settle the question of whether or not imageability has a role in picture naming.

Next, we conducted a multiple regression analysis on naming agreement (NA_{st}) following the approach of Snodgrass and Yuditsky (1996) and Cuetos and Alija (2003). Table 4 also shows these results. The predictor variables were IA, IV, IMG, VIV, FAM, AoA, CMP, LogFREQ, and MORA. In this analysis, H was not used as an independent variable, because NA_{st} and H were highly correlated ($-.88$). VIV strongly predicted naming agreement. Also AoA, FAM, and IA contributed moderately to the prediction of naming agreement. Lexical properties of the words—that is, LogFREQ and MORA—also made small contributions. None of the other independent variables were significant.

The results of the regression analyses revealed that, along with conventional independent variables such as name diversity and AoA, imagery-related factors such as VIV and IA were reliable predictors of naming time and agreement. This may suggest that certain processes involving mental imagery have a role in picture naming. However, previous studies have primarily focused on verbal aspects of picture naming, placing little emphasis on its nonverbal aspects, such as mental imagery. Future research into picture-naming processes should consider such image-related properties.

Conclusions

The present study had two aims. First, we investigated whether picture naming could be predicted by variables not sufficiently considered in picture processing, particularly imagery-related variables (imageability, vividness, image agreement, and image variability). Second, we provided refined Japanese normative data for 360 pictures based on Nishimoto et al. (2005). Naming times and name agreement were collected for each picture. All together, a total of 17 measures were obtained for each picture.

We found that the refined picture set improved significantly in name agreement as compared with the prior set. Almost all correlations among the measures were statistically significant. However, correlations on imageability were not significant, nor were most of the correlations involving complexity. A simultaneous multiple regression analysis was carried out on naming times, using a range of other measures as independent variables. The most reliable predictor was naming diversity, followed by image agreement, age of acquisition, and vividness. As for the four

imagery-related measures, two of these—namely, image agreement and vividness—contributed substantially to predictions of naming times. Furthermore, we found that naming accuracy (measured as name agreement) was predicted by vividness, followed by age of acquisition, familiarity, and image agreement. The results of regression analyses suggested that certain processes involving mental imagery have a role in picture naming.

References

- Alario, F.-X., & Ferrand, L. (1999). A set of 400 pictures standardized for French: Norms for name agreement, image agreement, familiarity, visual complexity, image variability, and age of acquisition. *Behavior Research Methods, Instruments, & Computers*, *31*, 531–552. doi:10.3758/BF03200732
- Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*, 140–155. doi:10.3758/BF03195559
- Amano, S., & Kondo, T. (2000). *Lexical properties of Japanese: Vol. 7. Frequency (in Japanese)*. Tokyo: Sanseido.
- Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency, and name agreement. *Quarterly Journal of Experimental Psychology*, *50A*, 560–585. doi:10.1080/027249897392026
- Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., . . . Tzeng, O. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*, 344–380. doi:10.3758/BF03196494
- Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language*, *50*, 456–476. doi:10.1016/j.jml.2004.02.001
- Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, *93*, 89–114. doi:10.1348/000712602162463
- Bonin, P., Peereman, R., Malardier, N., Méot, A., & Chalard, M. (2003). A new set of 299 pictures for psycholinguistic studies: French norms for name agreement, image agreement, conceptual familiarity, visual complexity, image variability, age of acquisition, and naming latencies. *Behavior Research Methods, Instruments, & Computers*, *35*, 158–167. doi:10.3758/BF03195507
- Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The bank of standardized stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS One*, *5*, e10773. doi:10.1371/journal.pone.0010773
- Carroll, J. B., & White, M. N. (1973a). Age of acquisition norms for 220 picturable nouns. *Journal of Verbal Learning and Verbal Behavior*, *12*, 563–576. doi:10.1016/S0022-5371(73)80036-2
- Carroll, J. B., & White, M. N. (1973b). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*, *25*, 85–95. doi:10.1080/14640747308400325
- Clark, J. M., & Paivio, A. (2004). Extensions of the Paivio, Yuille, and Madigan (1968) norms. *Behavior Research Methods, Instruments, & Computers*, *36*, 371–383. doi:10.3758/BF03195584
- Coltheart, M. (1981). The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, *33A*, 497–505. doi:10.1080/14640748108400805

- Cuetos, F., & Alija, M. (2003). Normative data and naming times for action pictures. *Behavior Research Methods, Instruments, & Computers*, 35, 168–177. doi:10.3758/BF03195508
- Cuetos, F., Ellis, A. W., & Alvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers*, 31, 650–658. doi:10.3758/BF03200741
- D'Angiulli, A. (2003–2004). Dissociating vividness and imageability. *Imagination, Cognition, and Personality*, 23, 79–88. doi:10.2190/JOG5-FTH8-8950-6Y8V
- D'Angiulli, A., & Reeves, A. (2007). The relationship between self-reported vividness and latency during mental size scaling of everyday items: Phenomenological evidence of different types of imagery. *The American Journal of Psychology*, 120, 521–551.
- de Groot, A. M. B. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15, 824–845. doi:10.1037/0278-7393.15.5.824
- Dell'Acqua, R., Lotto, L., & Job, R. (2000). Naming times and standardized norms for the Italian PD/DPSS set of 266 pictures: Direct comparisons with American, English, French, and Spanish published databases. *Behavior Research Methods, Instruments, & Computers*, 32, 588–615. doi:10.3758/BF03200832
- Desrochers, A., & Bergeron, M. (2000). Value of subjective frequency and imagery for a sample of 1,916 French language nouns. *Canadian Journal of Experimental Psychology*, 54, 274–325. doi:10.1016/S1196-1961(02)00073-9
- Desrochers, A., & Thompson, G. L. (2009). Subjective frequency and imageability ratings for 3,600 French nouns. *Behavior Research Methods*, 41, 546–557. doi:10.3758/BRM.41.2.546
- Dimitropoulou, M., Duñabeitia, J. A., Blitsas, P., & Carreiras, M. (2009). A standardized set of 260 pictures for modern Greek: Norms for name agreement, age of acquisition, and visual complexity. *Behavior Research Methods*, 41, 584–589. doi:10.3758/BRM.41.2.584
- Ellis, A. W., & Morrison, C. M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 515–523. doi:10.1037/0278-7393.24.2.515
- Frost, S. J., Mencl, W. E., Sandak, R., Moore, D. L., Rueckl, J. G., Katz, L., . . . Pugh, K. R. (2005). A functional magnetic resonance imaging study of the tradeoff between semantics and phonology in reading aloud. *NeuroReport*, 16, 621–624. doi:10.1097/00001756-200504250-00021
- Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, 5, 67–104. doi:10.1080/02643298808252927
- Izura, C., Hernández-Muñoz, N., & Ellis, A. (2005). Category norms for 500 Spanish words in five semantic categories. *Behavior Research Methods*, 37, 385–397. doi:10.3758/BF03192708
- Jefferies, E., Patterson, K., Jones, R. W., & Ralph, M. A. (2009). Comprehension of concrete and abstract words in semantic dementia. *Neuropsychology*, 23, 492–499. doi:10.1037/a0015452
- Johnston, R. A., Dent, K., Humphreys, G. W., & Barry, C. (2010). British-English norms and naming times for a set of 539 pictures: The role of age of acquisition. *Behavior Research Method*, 42, 461–469. doi:10.3758/BRM.42.2.461
- Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic assessments of language processing in aphasia*. Hove, U.K.: Erlbaum.
- Kosslyn, S. M. (1994). *Image and brain*. Cambridge, MA: MIT Press.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and Brain Sciences*, 22, 1–75. doi:10.1017/S0140525X99001776
- Lloyd-Jones, T. J., & Nettlemill, M. (2007). Sources of error in picture naming under time pressure. *Memory & Cognition*, 35, 816–836. doi:10.3758/BF03193317
- Lotto, L., Surian, L., & Job, R. (2010). Objective age of acquisition for 223 Italian words: Norms and effects on picture naming speed. *Behavior Research Methods*, 42, 126–133. doi:10.3758/BRM.42.1.126
- Manoiloff, L., Artstein, M., Canavoso, M. B., Fernández, L., & Segui, J. (2010). Expanded norms for 400 experimental pictures in an Argentinean Spanish-speaking population. *Behavior Research Methods*, 42, 452–460. doi:10.3758/BRM.42.2.452
- Matsukawa, J. (1983). A study of characteristics of pictorial material: I (in Japanese). *Memoirs of the Faculty of Law & Literature, Shimane University*, 6, 97–139.
- Morrison, C. M., Chappell, T. D., & Ellis, A. W. (1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology*, 50A, 528–559. doi:10.1080/027249897392017
- Morrison, C. M., Ellis, A. W., & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming, not object recognition. *Memory & Cognition*, 20, 705–714. doi:10.3758/BF03202720
- Nishimoto, T., & Hayashi, S. (1996). A standardized set of 143 pictures: Norms for name agreement, image agreement, and familiarity (in Japanese). *Waseda Psychological Reports*, 28, 59–85.
- Nishimoto, T., Miyawaki, K., Ueda, T., Une, Y., & Takahashi, M. (2005). Japanese normative set of 359 pictures. *Behavior Research Methods*, 37, 398–416. doi:10.3758/BF03192709
- Nishimoto, T., & Yasuda, Y. (1982). A standardized set of 260 pictures: Japanese norms for name agreement, image agreement, and visual complexity (in Japanese). *Waseda Psychological Reports*, 14, 55–76.
- Paivio, A., Clark, J., Digidon, N., & Bons, T. (1989). Referential processing: Reciprocity and correlates of naming and imaging. *Memory & Cognition*, 17, 163–174. doi:10.3758/BF03197066
- Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, 76(1, Pt. 2), 1–25. doi:10.1037/h0025327
- Pérez, M. A. (2007). Age of acquisition persists as the main factor in picture naming when cumulative word frequency and frequency trajectory are controlled. *Quarterly Journal of Experimental Psychology*, 60, 32–42. doi:10.1080/17470210600577423
- Pind, J., Jónsdóttir, H., Gissurardóttir, H., & Jónsson, F. (2000). Icelandic norms for the Snodgrass and Vanderwart (1980) pictures: Name and image agreement, familiarity, and age of acquisition. *Scandinavian Journal of Psychology*, 41, 41–48. doi:10.1111/1467-9450.00169
- Pind, J., & Tryggvadóttir, H. B. (2002). Determinants of picture naming times in Icelandic. *Scandinavian Journal of Psychology*, 43, 221–226. doi:10.1111/1467-9450.00290
- Sakuma, N., Ijuin, M., Fushimi, T., Tatsumi, I. F., & Tanaka, M. (2005). *Nihongo-no goitokusei* [Lexical properties of Japanese]: Vol. 8. Word imageability (in Japanese). Tokyo: Sanseido.
- Sanfeliu, M. C., & Fernandez, A. (1996). A set of 254 Snodgrass–Vanderwart pictures standardized for Spanish: Norms for name agreement, image agreement, familiarity, and visual complexity. *Behavior Research Methods*, 28, 537–555. doi:10.3758/BF03200541
- Severens, E., Van Lommel, S., Ratinckx, E., & Hartsuiker, R. J. (2005). Timed picture naming norms for 590 pictures in Dutch. *Acta Psychologica*, 119, 159–187. doi:10.1016/j.actpsy.2005.01.002
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215. doi:10.1037/0278-7393.6.2.174
- Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, 28, 516–536. doi:10.1037/0278-7393.6.2.174

- Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1140–1154. doi:10.1037/0278-7393.21.5.1140
- Székely, A., & Bates, E. (2000). Objective visual complexity as a variable in studies of picture naming. *Center for Research in Language Newsletter*, *12*. La Jolla, CA: University of California, San Diego. Available at www.crl.ucsd.edu/newsletter/12-2/article.html
- Székely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., . . . Bates, E. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research Methods, Instruments, & Computers*, *35*, 621–633. doi:10.3758/BF03195542
- van Hell, H. J. G., & de Groot, A. M. B. (1998). Disentangling context availability and concreteness in lexical decision and word translation. *Quarterly Journal of Experimental Psychology*, *51A*, 41–63. doi:10.1080/027249898391756
- van Loon-Vervoorn, W. A. (1985). *Voorstelbaarheidswaarden van Nederlandse Woorden: 4600 substantieven, 1000 verba, 500 adjectieven* [Imageability ratings of Dutch words: 4,600 nouns, 1,000 verbs, 500 adjectives]. Lisse, Switzerland: Swets & Zeitlinger.
- Viggiano, M. P., Vannucci, M., & Righi, S. (2004). A new standardized set of ecological pictures for experimental and clinical research on visual object processing. *Cortex*, *40*, 491–509. doi:10.1016/S0010-9452(08)70142-4
- Vitkovitch, M., & Tyrrell, L. (1995). Sources of disagreement in object naming. *Quarterly Journal of Experimental Psychology*, *48A*, 822–848. doi:10.1080/14640749508401419
- Weekes, B. S., Shu, H., Hao, M., Liu, Y., & Tan, L. H. (2007). Predictors of timed picture naming in Chinese. *Behavior Research Methods*, *39*, 335–342. doi:10.3758/BF03193165