University of Reading

What does a cue do? Comparing phonological and semantic cues for picture naming in aphasia

Article

Accepted Version

Meteyard, L. and Bose, A. (2018) What does a cue do? Comparing phonological and semantic cues for picture naming in aphasia. Journal of Speech, Language, and Hearing Research, 61 (3). pp. 658-674. ISSN 1558-9102 doi: https://doi.org/10.1044/2017_JSLHR-L-17-0214 Available at http://centaur.reading.ac.uk/75215/

To link to this article DOI: http://dx.doi.org/10.1044/2017_JSLHR-L-17-0214

Publisher: ASHA

JSLHR JOURNAL OF SPEECH, LANGUAGE, AND HEARING RESEARCH

http://jslhr.pubs.asha.org/journal.aspx

Online ISSN: 1558-9102

Print ISSN: 1092-4388

What does a cue do? Comparing phonological and semantic cues for picture naming in aphasia

Lotte Meteyard

Arpita Bose

To cite this article: Meteryard, L. & Bose, A. (in-press). What does a cue do? Comparing phonological and semantic cues for picture naming in aphasia. Journal of Speech Language and Hearing Research.

Publisher: American Speech and Hearing Association

Journal: Journal of Speech, Language, and Hearing Research (JSHLR)

Running Head: Cueing in aphasia

What does a cue do? Comparing phonological and semantic cues for picture naming in aphasia

Lotte Meteyard

Arpita Bose

Department of Clinical Language Sciences

University of Reading

United Kingdom

Keywords: Cueing, phonology, semantic, naming, lexical variables.

Corresponding author:
Arpita Bose, Ph.D.
Department of Clinical Language Sciences,
School of Psychology and Clinical Language Sciences,
University of Reading, Reading, UK.
Email: a.bose@reading.ac.uk

Abstract

Purpose: Impaired naming is one of the most common symptoms in aphasia, often treated with cued picture naming paradigms. It has been argued that semantic cues facilitate the reliable categorisation of the picture, and phonological cues facilitate the retrieval of target phonology. To test these hypotheses, we compared the effectiveness of phonological and semantic cues in picture naming for a group of individuals with aphasia. To establish the locus of effective cueing, we also tested whether cue type interacted with lexical and image properties of the targets.

Method: Individuals with aphasia (n=10) were tested with a within-subject design. They named a large set of items (n=175) four times. Each presentation of the items was accompanied by a different cueing condition (phonological, semantic, non-associated word and tone). Item level variables for the targets (i.e., phoneme length, frequency, imageability, name agreement and visual complexity) were used to test the interaction of cue type and item variables. Naming accuracy data was analysed using generalised linear mixed effects models.

Results: Phonological cues were more effective than semantic cues, improving accuracy across individuals. However, phonological cues did not interact with phonological or lexical aspects of the picture names (e.g., phoneme length, frequency). Instead, they interacted with properties of the picture itself (i.e., visual complexity), such that phonological cues improved naming accuracy for items with low visual complexity.

Conclusions: The findings challenge the theoretical assumptions that phonological cues map to phonological processes. Instead, phonological information benefits the earliest stages of picture recognition, aiding the initial categorization of the target. The data help to explain why patterns of cueing are not consistent in aphasia, i.e., it is not the case that phonological impairments always benefit from phonological cues and semantic impairments form semantic cues. A substantial amount of the literature in naming therapy focuses on picture naming paradigms. Therefore, the results are also critically important for rehabilitation, allowing for therapy development to be more rooted in the true mechanisms through which cues are processed.

Introduction

Impaired naming is the most common symptom of word production difficulties for people with aphasia (PWA). Cueing is a ubiquitous technique used both in assessment and therapy to ameliorate naming impairments (Laine & Martin, 2006). A cue is a piece of relevant linguistic information presented once, prior to the individual attempting to name the target or after a failed production attempt. Typical cues are phonological (the first sound, e.g. “k” for *cup*) or semantic (a related word, e.g. “purr” for *cat*) (Heath et al., 2012; Nickels, 2002; Li & Williams, 1990). If the cue is effective, it will facilitate word production and result in more accurate naming (Nickels & Best, 1996). In assessment, cueing has been used to establish the nature of the naming impairments and when applied systematically over a long period of time and often as a hierarchy (e.g. first sound, first syllable, whole word), cueing becomes a therapeutic intervention (Nickels & Best, 1996). Numerous therapies for naming utilize cues in the context of picture naming (e.g., Best, Greenwood, Grassly, Herbert, Hickin & Howard, 2013; Kiran & Bassetto, 2008; Leonard, Rochon & Laird, 2008; Van Hees, Angwin, McMohan & Copland, 2013).Despite the long history of use, there is no clear understanding of how, specifically, cues improve naming performance for PWA (Heath et al., 2012; Lorenz & Nickels, 2007; Pellet Cheneval, Bonnans, & Laganaro, 2017).

Irrespective of theoretical models, it is widely accepted that naming an object involves at least two stages: a) retrieval of the semantic information, and b) attaching form to the selected word[1] (Schwartz, 2013). A simple hypothesis is that phonological and semantic cues support the retrieval of phonological and semantic information for a target word,

[1] Beyond this basic framework there are many differences amongst the current models, with disagreements regarding the specifics within semantic and phonological levels, as well as discreteness and interactivity between these stages (Foygel & Dell, 2000; Levelt, Roelofs, & Meyer, 1999; Rapp & Goldrick, 2000). For example, whilst Dell and colleagues (Dell et al., 1997) generally refer to lexical units as word units, the term is often used interchangeably with the ‘lemma’ of Levelt and colleagues’ model (Levelt et al., 1999). There is also debate concerning the existence of intermediary representation between semantics and phonology (e.g. Caramazza, 1997). Within this class of models, some propose the discreetness of the stages (Levelt et al., 1999), whilst others postulate an interactive flow of information between them (Foygel & Dell, 2000; Rapp & Goldrick, 2000).

respectively (e.g., Cuetos, Aguado, & Caramazza, 2000; Howard & Gatehouse, 2006; Jefferies & Lambon Ralph, 2006). Similarly, a linked hypothesis is that phonological cues should remediate phonological impairments and semantic cues should remediate semantic impairments (e.g., Hickin, Best, Herbert, Howard, & Osborne, 2002; Van Hees, Angwin, McMahon, & Copland, 2013). Although it has been shown that individuals who experience successful facilitation with cueing during assessment also respond to cue based therapy (Hickin et al., 2002), there are few reliable correlations between individual patient profiles and response to cueing therapy (Lorenz & Ziegler, 2009; Wisenburn & Mahoney, 2009). This may be due to a combination of heterogeneous aphasic profiles and small neuropsychological sample sizes. To mitigate these difficulties, a few studies have used within-subject designs (e.g., Davis & Pring, 1991; Lorenz & Ziegler, 2009; Van Hees et al., 2013).

The provision of cues during picture naming has been compared to repetition priming (Martin & Laine, 2000). That is, cues provide a short-term benefit from repeated presentations with a specific target item and the association this creates in a particular task (Logan, 1990). Beyond immediate facilitation, cues can lead to improved naming of target items at longer lags, e.g., more than 6 intervening items, 10 minutes (Heath et al., 2012; 2013). There is now a consensus that phonological and semantic therapies are equally effective (Lorenz & Ziegler, 2009); the difference between them is often 'overstated' (Davis & Pring, 1991). This is likely because word phonology is activated in semantic tasks and semantics is activated in phonological tasks, especially in picture naming where the picture stimulus activates conceptual and semantic information (e.g., Howard, Hickin, Redmond, Clark & Best, 2006).

There was early evidence that phonological cues help to specify a semantic target for the picture, which in turn aids the selection of a specific phonological form (Stimely & Noll, 1991; Li & Williams, 1991). In healthy adults, the presence of linguistic information (e.g., a verbal label) improves categorization and learning of visual stimuli (Lupyan, 2008). Thus, in both healthy and impaired language processing, exposure to target related phonology improves categorization of target stimuli such as pictures. In the few studies that have directly compared phonological and semantic cues (i.e. within-subjects), some common patterns emerge. Whilst both phonological and semantic cues can be effective (sometimes equally across a group, Stimely & Noll, 1991), phonological cues tend to be effective for more individuals (Van Hees et al., 2013; Lorenz & Ziegler, 2009; Li & Williams, 1991). It appears likely that phonological cues facilitate both the visual categorization of the picture as well as priming output phonology, making them more useful in picture naming than semantic cues. For semantic cues, pre-exposure to the picture alongside a semantic task (i.e., property-picture verification such as verbal presentation "Does it purr?" with a picture of a cat) provided both short and long-term facilitation of picture naming (Heath et al., 2012; 2013). Under this description, perceptual and/or lexico-semantic processing of the picture is made more efficient by pre-exposure, and this improves naming.

To recap, naming a picture means to establish a reliable categorization of the picture (i.e. stable semantic information) and produce the associated target specific output phonology. Evidence suggests that semantic as well as phonological cues improve picture naming. However, the precise mechanism remains unclear. Yet the literature points towards an explanation where cues make the form-meaning mapping more reliable, rather than selectively improving a phonological or semantic stage of processing. For example, cutting across classical aphasia categories, it has been shown that individuals with better semantic processing and worse output phonology benefit from naming therapy (Best et al., 2013; Howard et al., 2006). These are precisely the individuals that can make use of cue information to improve the mapping from form to meaning and (re)learn target phonology.

Our attempts to improve word production can be substantially improved by understanding why and how cues facilitate naming in aphasia. A fruitful way to address this question may be to use variation in the lexical properties of items. For example, words differ along a of range properties (e.g., word frequency, imageability, length) as do pictures (e.g., visual complexity, name agreement). A large body of psycho- and neuro-linguistic have shown the effect of specific lexical and image properties on word production (e.g., Indefrey, 2011). In healthy adults as well as in PWA, the general findings in the literature is that items are named more quickly and accurately when they are shorter, more frequent, more imageable and have high name agreement (e.g., Bose & Schafer, 2017; Alario et al., 2004; Bonin, Peereman, Malardier, Méot, & Chalard, 2003; Kittredge, Dell, Verkuilen & Schwartz, 2008; Nickels & Howard, 1995; Middleton & Schwartz, 2010; Vitevitch & Sommers, 2003). In addition, a retrospective analysis of items from three therapy studies found that highly imeagable words were named more accurately and required less cueing (Conroy, Snell, Sage & Lambon Ralph, 2012).

To the best of our knowledge no study has explored the interaction between cue type (phonological or semantic) and target characteristics for PWA. Therefore, the aim of this study was to investigate the effect of cue condition (phonological vs. semantic) on the picture naming accuracy for a large set of items (175 pictures) for a mixed group of ten PWA. For additional experimental control and to aid interpretation, the study also included a neutral cue condition, to allow the effect of phonological and semantic cues to be properly compared and evaluated (Stimely & Noll, 1991). The study also explored how cues interact with characteristics of the target words. Mixed model analysis allowed us to consider multiple predictor variables (condition, cue, and lexical variables) and account for variation in participant performance across conditions. Our research questions were:

1. Are phonological and semantic cues equally effective in improving naming accuracy?

2. Is there an interaction between the cue type (phonological / semantic) and the lexical and image properties of the word?

If there is a simple mapping between cue type (phonological / semantic) and target properties, we should see the following: 2a. Semantic cues should facilitate the retrieval of conceptual and semantic information. We should see an interaction between semantic cues and imageability (i.e. words with lower imageability should benefit more from a semantic cue) and with name agreement (i.e. words with lower name agreement should benefit more from a semantic cue, e.g. Bose & Schafer, 2017). 2b. Semantic and phonological cues should facilitate the retrieval of lexical information. We should see an interaction of semantic and phonological cues with frequency (i.e. words with lower frequency should benefit from both semantic and phonological cues). 2c. Phonological cues should facilitate the retrieval of phonological information. We should see an interaction between phonological cues and length (i.e. words that are longer should benefit from phonological cues).

Method

Participant sample and Background Test Battery

We recruited a mixed group of ten participants with aphasia (PWA, 6 male, 4 female). Age ranged from 42 to 85 years (*M*=68.8, *SD*=13.75), education level from 13 to 16 years (*M*=14.7, *SD*=1.16), and they were ten to 168 months (*M*=64.3, *SD*=54.7) post-onset of stroke. Inclusion criteria for PWA were: a single left hemisphere cardiovascular accident as determined by neuroradiological and/ neurological examinations; a diagnosis of aphasia on standardized clinical tests (Boston Diagnostic Aphasia Examination, Goodglass, Kaplan, & Barresi, 2001); at least eight months post-stroke; monolingual English speaker; no history of other neurological illness, psychiatric disorders or substance abuse; no visual field or sensory perceptual deficits based on the Reitan-Klove Sensory Perceptual Examination (Reitan & Wolfson, 1993); and no other significant cognitive deficits.

For sample size calculation, power analysis was completed using data from Lorenz & Ziegler (2009) that compared different cue types within-participants for a group of PWA, from which we could extract data for each individual to calculate mean, standard deviation and correlations (i.e. the correlation of participants' scores in the two conditions). We used equation 8 from Morris & DeShon (2002) which takes the difference between the two condition means, and divides by the average standard deviation multiplied by a product of the correlation between conditions. This equation means that the resultant effect size can be interpreted like the more familiar between-subjects effect size calculation (see Morris & DeShon, 2002, for details). Lorenz & Ziegler (2009) found a benefit of phonological cueing within subjects (no cue mean proportion correct M = 65.80, SD = 46.09, cued additional proportion correct M = 78.40, SD = 50.72, correlation = 0.98) that gave an effect size of 1.30. Within subject differences between proportion correct for word-form and semantic cues gave an effect size of 1.63 (word form M = 29.30, SD = 11.97; semantic M = 8.80, SD = 4.29, correlation = -0.20). When comparing cue conditions within subjects, a power of 0.95 would be achieved with samples of 8 or 6 respectively. Therefore, our sample should be sufficient to detect effects of cueing (i.e. when comparing phonological cues to semantic cues, or phonological cues to a control/absent cue condition). We were unable to complete an a-priori power analysis for the interaction between cue type and target variables, and the mixed effects analysis that includes the variances associated with the random effects (the groupings of participants and items) as data was not available for this. However, we anticipated that the number of items (175) presented to each participant for naming in each cue condition would be sufficient to detect main effects of relevant lexical variables. Given the large effect sizes of cueing from Lorenz & Ziegler (2009) we were hopeful then that interactions between cue type and target variables could be detected. In addition, by manipulating cue type and target variables within subjects we were hopeful this would increase sensitivity to any interactions.

A comprehensive evaluation of participants' single word comprehension and production was performed using subtests from the Psycholinguistic Assessments of Language Processing in Aphasia (PALPA, Kay, Lesser, & Coltheart, 1992); the 3-picture version of the Pyramids and Palm Trees Test (PPT, Howard & Patterson, 1992) and the Philadelphia Naming Test (PNT, Roach, Schwartz, Martin, Grewal, & Brecher, 1996). This battery measured overall picture naming abilities, input and output phonological abilities, and conceptual and lexico-semantic processing. Table 1 presents demographic information, aphasia type, severity and results of the background assessments for each of the PWA. Written informed consent procedures in accordance to the University Research Ethics Board were followed for all participants.

The group included four individuals with Broca's aphasia, two with transcortical motor aphasia, one with mixed aphasia, two with anomic aphasia, and one with Werincke's aphasia. BDAE aphasia severity ratings ranged from 1 to 3.5 (*M*=1.75, *SD*=0.89; 1 as most severe and 5 the least severe). They showed a wide range in picture naming abilities (PNT scores ranged from 20% to 87%, *Mean*= 49.8, *SD*= 23.6). As a group they showed variable impairments both for input and output phonology but with better preserved conceptual and lexical semantics (PPT scores ranged from 83% to 98%, Mean = 91.7, SD = 5.0). See Table 1.

Table 1. Demographic details, aphasia type and severity, and performance (% correct) on language tasks for People with Aphasia (PWA)

Participants	JV	JK	DH	WR	EW	FF	MH	AM	AW	CB	*Mean*	*SD*
Age (years)	78	62	56	42	75	76	73	85	83	58	68.80	13.75
Sex	F	M	F	M	M	M	F	M	F	M		
Education (years)	13	14	15	16	14	15	16	16	15	13	14.70	1.16
Months post-stroke	60	149	42	16	10	30	60	168	24	84	64.30	54.70
Aphasia type	Anomic	Broca's	Anomic	Broca's	Broca's	Wernicke's	TCM	Mixed	Broca's	TCM		
BDAE severity	3	1	3.5	1	1.5	1.5	1	1	2	2	1.75	0.89
Naming and word production across modalities												
Philadelphia Naming Test [1]	72	33.1	39	54.3	52.6	34	25	20	80.6	87	49.8	23.6
Number of Correct (N=175)	126	58	68	95	92	60	44	35	141	150	86.9	40.8
Error profile (#, proportion)												
Formal errors	10 (0.20)	4 (0.03)	8 (0.07)	3 (0.04)	0 (0.00)	4 (0.03)	3 (0.02)	8 (0.06)	2 (0.06)	0 (0.00)		
Semantic errors	6 (0.12)	46 (0.39)	24 (0.22)	33 (0.41)	10 (0.12)	18 (0.16)	25 (0.19)	11 (0.08)	13 (0.38)	17 (0.68)		
Mixed errors	1 (0.02)	5 (0.04)	4 (0.04)	4 (0.05)	5 (0.06)	0 (0.00)	0 (0.00)	0 (0.00)	3 (0.09)	0 (0.00)		
Nonword errors	10 (0.20)	5 (0.04)	3 (0.03)	5 (0.06)	3 (0.04)	57 (0.50)	0 (0.00)	58 (0.41)	5 (0.15)	2 (0.08)		
Unrelated errors	0 (0.00)	14 (0.12)	5 (0.03)	1 (0.01)	8 (0.10)	8 (0.07)	0 (0.00)	31 (0.22)	4 (0.12)	1 (0.04)		
Miscellaneous	22 (0.45)	43 (0.37)	63 (0.59)	34 (0.43)	57 (0.69)	28 (0.24)	103 (0.59)	32 (0.23)	7 (0.21)	5 (0.2)		
Across Modality (#53 PALPA[2])												
#53 Repetition	92.5	92.5	97.5	100	92.5	95	88	65	95	97.5	91.6	9.9
#53 Naming	75	30	42.5	55	62.5	40	30	25	85	95	54.0	24.7
Input and output phonology												
#2 PALPA: Real word minimal pair discrimination	88.88	76.39	86.11	90.27	94.4	92	92	62.5	75	97.2	85.5	10.8
Same	97.22	94.44	97.22	97.22	100	100	97	77.8	81	97.2	93.9	7.8
Different	80.55	58.33	72.22	83.33	88.8	83.3	86	47.2	69	97.2	76.6	15.1
#4 PALPA: Minimal pair requiring picture selection	100	95	87.5	97.5	72.5	88	93	77.5	82.5	100	89.4	9.5
# 9 PALPA: Word repetition	85	66.25	95	98.75	94	96.3	100	20	62.5	100	81.8	25.7
High imageability	92.5	77.5	95	100	97.5	97.5	100	35	75	100	87.0	20.4
Low imageability	77.5	55	95	97.5	95	95	100	5	50	100	77.0	31.4
High frequency	90	72.5	97.5	100	97.5	97.5	100	20	52.5	100	82.8	27.0

Low frequency	80	60	92.5	97.5	90	95	100	20	72.5	100	80.8	25.0
# 8 PALPA: Nonword repetition	23.33	53.33	76.67	76.66	66.6	87	57	3.3	40	83.3	56.7	27.4
1-syllable	30	40	80	70	40	90	50	10	10	80	50.0	29.1
2-syllable	30	50	60	80	80	80	70	0	60	80	59.0	26.4
3-syllable	10	70	90	80	80	90	50	0	50	90	61.0	33.1
Conceptual and lexico-semantic processing												
Pyramids and Palm Trees[3]	96.15	82.69	98.08	92.31	86.5	89	90	94.2	90.4	98	91.7	5.0
# 47 PALPA: Spoken word-picture matching	100	97.5	95	92.5	85	93	85	100	92.5	95	93.6	5.3
#49 PALPA: Auditory synonym judgments	85	70	88.33	76.66	81.6	86.7	86.7	53.3	85	88.3	80.2	11.1
High imageability	93.33	80	96.67	80	99.7	90	96.7	63.3	96.7	100	89.6	11.8
Low imageability	76.66	60	80	73.33	66.7	83.3	76.7	43.3	73.3	76.7	71.0	11.8

[1] Philadelphia Naming Test (Roach et al., 1996); [2]Psycholinguistic Assessments of Language Processing in Aphasia (Kay et al., 1992); [3] Pyramid and Palm Trees Test (Howard & Patterson, 1989); [4] Transcortical Motor Aphasia.

Experimental manipulation

A four-session computerized picture naming experiment was developed using the 175-item PNT (Roach et al., 1996). Four testing sessions manipulated cue condition (semantic vs. phonological) and cue type (valid phonological cue vs. control tone; valid semantically related cue vs. non-associated control). The sessions were blocked by condition (semantic and phonological): two sessions presented semantic cues and two sessions presented phonological cues. For each condition, testing was conducted over two sessions such that the items that were preceded by valid cues in one session were preceded by the control cues in the other session and vice versa. For the phonological cueing sessions, the auditory cue was either the first sound of the name of the picture (e.g., /ball/→ "b") in the valid condition or 1KHz pure-tone in the control condition. For semantic cueing, the valid auditory cue was a semantically related word (e.g., /candle/→"wick") or a semantically non-associated word in the control condition (e.g., /candle/→"chop"). Experimental and therapeutic studies have used wide range of semantic cues (e.g., semantic attributes, concept properties, word associates, category memberships, etc). To generate our semantic cues, we used the first associates of the target items in the University of South Florida Word Association Norms (Nelson, McEvoy, & Schreiber, 1998). This allowed a quantified measure for the semantic relationship, rather than subjective or intuited methods of generating semantic cues. Associated cues have also been shown to affect picture naming responses. If the first associate had the same phoneme onset with the target, we choose the next associate to ensure that there was no phoneme overlap with the cue and the target. Therefore, none of the semantic cues (associated or non-associated controls) had the same initial phoneme as the target. The non-associated control semantic cues were also matched to the related cues for word frequency, syllable length, and familiarity, and the non-associated semantic cues did not appear in the possible words in the association norms for that target word. The verbal cues were generated by a native English-speaking female and recorded in a sound-attenuated room. The pure tone was computer generated. The stimuli items are freely available on the Moss Aphasia Psycholinguistic Project Database (http://www.mappd.org/). Appendix A provides the stimuli with their lexical properties and the cues used in this experiment. A copy of the experiment software is available upon request.

A trial consisted of the presentation of the recorded auditory cue, followed by 750ms of silence and then the target picture, which remained on the computer screen until a response was made or a maximum of 10 seconds had lapsed. A delay of 750ms between cue and target was selected on the basis of a review of the aphasia picture naming literature. This literature shows delays ranging from 350 ms to 1400ms (Baum, 1997; Hagoort, 1997). Facilitation from cueing in picture naming has been observed for both positive and negative stimulus onset asynchronies (i.e. when the cue is presented both before or after the picture). However, a greater number of experimental studies have chosen a positive stimulus onset asynchrony (e.g., Indefrey, 2011) and demonstrated facilitation in picture naming from the cue being presented before the picture (e.g. Bose & Buchanan, 2007). For that reason, we selected a duration of 750ms that would allow sufficient time the cue to be processed but would not unnecessarily lengthen each trial (as participants were already completing four sessions of 175 naming trials). Other than occasional encouragement during the session, no other feedback was provided to the participants. The stimuli presentation was randomized within sessions, and validness and cue condition were counterbalanced across sessions. Therefore, for each participant there were a total of 700 picture naming trials [175 items X 2 cue condition (semantic and phonological) X 2 cue types (valid vs. control)]. There was a gap of at least one week between each of the four sessions. These sessions were recorded with a high quality digital audio recorder and later transcribed for analysis.

Scoring and reliability

We followed Roach et al. (1996) criteria to score the naming responses. The first complete non-fragmented naming attempt was scored as correct or incorrect. Criteria for identifying the first complete response were: ignoring instances of single phonemes or consonant + schwa and filler items (e.g. um, uh) altogether, incomplete items are judged on the basis of auditory cues (e.g. segment duration, a lack of downward or questioning intonation, no pause separating an item from the following attempt) which indicate self-interruption. Responses were scored as correct if they replicated the target name. Addition or deletion of plural morphemes was accepted, as was the addition of modifiers such as "wishing well" for "well". All scoring was performed by the second author and a trained research assistant performed reliability checks for 35% of the sessions. The point-by-point inter-rater agreement was 96% (Cohen's kappa, $\kappa = 0.92$), and disagreements were resolved by reviewing the scoring definitions and the transcripts (4% of the data, 98 trials).

Influence of word properties

We selected target properties known to be influential in picture-naming: length (in phonemes), frequency (lemma), imageability, name agreement and visual complexity. Values for five target properties were obtained for each of the 175 PNT items (Table 2). Length (number of phonemes) is provided with the PNT materials. Log-transformed word frequency (lemma, per million) was retrieved from the CELEX lexical databases (Baayen, Piepenbrock, & van Rijn, 1993). A full set of imageability ratings for the PNT stimuli were not available, so 38 healthy controls provided ratings for all 175 pictures, with the subsequent set of ratings consisting of the mean across all participants for each item. Name agreement values consisted of the mean accuracy of 20 control participants whose PNT naming results are provided in the Moss Aphasia Psycholinguistics Project Database (http://www.mappd.org/). Visual complexity values (based on image file size) were obtained from Székely & Bates (2000). All properties were mean centered and scaled as z-scores to reduce collinearity.

Table 2. Descriptive statistics for lexical and image properties for the Philadelphia Naming Test items.

Target property	Mean	Range	SD
Length in phonemes	4.50	1.0-11.0	1.77
Log Lemma Frequency (per million)	1.34	0.00-3.21	0.63
Imageability	6.15	3.95-6.84	0.42
Name Agreement	0.98	0.75-1.00	0.05
Visual Complexity	217.50	59.00-526.00	89.16

Analysis

Generalised linear mixed effects models were used to model the data, implemented in R (R Core team, 2013) using the package lme4 (Bates, Maechler, Bolker & Walker, 2014), lmerTest (Kuznetsova, Brockhoff & Christensen, 2016) and effects (Fox, 2003). Accuracy was the dependent variable (binomial link function), giving log odds of producing a correct response as the model outcome. Random effects were used to model the experiment structure. Confidence intervals (95% CI) were calculated using the Wald method. We fit random intercepts and correlated slopes for cue type varying across participants and items. Intercepts and slopes for items were perfectly correlated and slopes did not improve model fit relative to intercepts only (X^2 = 0.3661, df = 2, p = 0.83); thus we retained only random intercepts for items. Fixed effects for session and trial were not significant and did not improve model fit (LogLik with Session = -3259, X^2 (1) = 0, p>.25; LogLik with Trial = -3259, X^2(1) = 0.64, p>.25). These were not included in further analysis. Models with fixed effects to test for cue, item properties and interactions of cue with item properties and Likelihood ratio tests comparing model fits are detailed below.

Results

Are phonological and semantic cues equally effective in improving naming accuracy?

We evaluated whether cue type improved naming performance. Cue type was entered as a fixed effect with four levels: phonological cue, tone, semantic cue and non-associated control word. This significantly improved model fit over the null model (vs. random effects only; LogLik with Cue = -3231, $X^2(3) = 54.5$, $p<.001$). The phonological cue condition significantly improved naming accuracy when compared against all other cues (all estimates >1.6; all 95% CIs between 0.37 – 3.00, see Table 3 for multiple comparisons and Figure 1), no other cue types affected naming performance.

Is there an interaction between the cue type and the item properties of the word?

The phonological cue was the only cue type that affected naming performance, with other cue types showing equivalent performance to each other. To simplify further analysis we collapsed across the cue conditions that were equal (semantic, non-associated and tone) to create a cue factor with two levels, phonological cue vs. no phonological cue (i.e. all other cue conditions). Modeling cue as a factor with two levels did not differ to a model with cue type as four levels (LogLik with Cue as two levels = -3232, X^2 (2) = 0.11, $p>.250$) and provided a significantly better fit to the data than a model with random effects only (LogLik Random Effects model = -3259, LogLik with Cue as two levels = -3232; $X^2(1) = 54.38$, $p<.001$). We evaluated whether item properties affected naming accuracy by entering each property as a main effect. Each item variable showed a significant main effect on naming accuracy and a model including main effects for all item properties was a significantly better fit than a model with cue type alone (LogLik with Cue alone = -3232, LogLik with Item main effects = -3179, $X^2(5) = 104.23$, $p<.001$). We tested the two-way interaction between each item property and cue, by adding interaction terms to a model with item properties as main effects[2].

Table 4 provides a summary of the final model. Items that were longer were named less accurately (slope estimate = -0.39, 95% CI = -0.54 - -0.25), items that were higher in imageability (estimate = 0.26, 95% CI = 0.11 - 0.41) and name agreement (estimate = 0.15, 95% CI = 0.04 - 0.26) were named more accurately (see Table 4 and Figure 2). Frequency did not significantly affect naming accuracy (estimate = 0.16, 95% CI = -0.02 - 0.33). The only item property to significantly interact with cue was visual complexity. The interaction between phonological cue and visual complexity gave a small improvement in model fit (LogLik with Item main effects = -3179, LogLik with Cue x Visual Complexity = -3178, X^2 (1) = 3.44, p=.064). When a phonological cue was present, visual complexity influenced naming accuracy (slope estimate = -0.25, 95% CI = -0.45 - -0.05). When a phonological cue was not present, there was no effect of visual complexity (estimate = -0.11, 95% CI = -0.25 - 0.02). Figure 2 shows that items with higher visual complexity were named with similar accuracy whether a phonological cue was present or not. For items of lower visual complexity, the phonological cue causes these items to be named more accurately.

Table 5 provides a summary of the odds ratios for each participant for each predictor in the model. These odds ratios show us the likelihood of a correct response without the influence of any predictors (the odds ratio for the intercept), and the odds ratios for each

[2] Following reviewer comments, we added a more complex random effects structure, entering random intercepts and correlated slopes for each item property varying over participants. This should better control for false positives in the model (Barr et al., 2013). When random intercepts and slopes were included the model showed signs of over-fitting, with perfect correlations between the intercept and slope variances for all item properties. This model also did not converge when random intercepts and correlated slopes were fit for imageability. We simplified by removing the intercepts (i.e. fitting slopes only) and testing against the maximal model (see Appendix C), until we achieved a simpler model that did not differ from the maximal model for goodness of fit. We also removed the random slope for Name Agreement as it had a variance of zero (i.e. did not explain any variance in the data); this model did not differ from the more complex random effects model (X^2 (9) = 13.768, p=0.13). We report the results from this model in Table 4, and results from the more complex random effects model in Appendix C (note that the significance of fixed effect predictors do not differ between the models). To check this model fit to the data, we refit the model after removing data points for which standardized residuals were greater than +/- 2.5 (38 data points). The results for the model were the same, except that the main effect of Frequency was significant. This model is reported in Appendix B.

predictor then tell us what increase or decrease in the odds of a correct naming response are present when we include specific predictors (e.g. a unit increase in length of the word, the presence of a phonological cue). Looking at the intercept values in Table 5, we see the variation in the naming ability for each participant (e.g. low odds of a correct response for participants AM, MH and FF; high odds of a correct response for participants CB, JV and EW). Turning to the predictors, across participants the presence of a phonological cue almost doubles the likelihood of getting a correct naming response (1.92). An increase in length (i.e. one more phoneme) reduces the likelihood of a correct naming response by around a third (~ -0.4). An increase in one unit of imageability increases the likelihood of a correct response by around a quarter (~0.25). For the interaction of visual complexity and the phonological cue, when a phonological cue is present an increase in visual complexity reduces the likelihood of a correct response by around a quarter (~ -0.25). As we saw in Figure 2, this is because items with low visual complexity benefit from a phonological cue whereas items with high visual complexity do not.

Table 3. Generalised linear mixed effects model results with multiple comparisons for cue type

Key: For each Cue comparison, the reference condition is the one to the right of the 'vs'.

Fixed Effects	**Estimate**	**Standard Error**	**95% CI**	**Z**	***p***[1]
Intercept	1.02	0.45	-0.08 - 2.11	2.27	0.023*
Phonological Cue vs Semantic Cue	1.69	0.50	0.37 - 3.00	3.38	0.004*
Phonological Cue vs Tone Control	1.65	0.47	0.42 - 2.87	3.54	0.002*
Phonological Cue vs Non-associated Control	1.69	0.50	0.37 - 3.00	3.38	0.004*
Semantic Cue vs Tone Control	-0.04	0.14	-0.41 - 0.32	-0.32	1
Semantic Cue vs Non-associated Control	0.00	0.09	-0.23 - 0.23	-0.01	1
Non-associated Cue vs Tone Control	0.04	0.14	-0.32 - 0.41	0.31	1
Random Effects	**Variance**	**Correlation**			
Items (intercept)	0.68				
Participants (intercept)	1.91				
Participant x Condition (slope)	0.11	-0.24			

[1]Bonferroni adjusted p-values. R model equation: Accuracy ~ (1+Condition|Subject) + (1|Word) + Cue Type

Table 4. Generalised linear mixed effects model results for phonological cue, target properties and their interactions.

Fixed Effects	Estimate	Standard Error	Wald 95% CI	Z	*p*
Intercept	0.31	0.44	-0.55 - 1.18	0.72	0.47
Phonological Cue vs No Phonological Cue	0.65	0.09	0.48 - 0.82	7.50	p<0.001*
Length	-0.39	0.07	-0.54 - -0.25	-5.32	p<0.001*
Frequency	0.16	0.09	-0.02 - 0.33	1.75	p=0.08
Imageability	0.26	0.08	0.11 - 0.41	3.42	p<0.001*
Name Agreement	0.15	0.0-6	0.04 - 0.26	2.74	p=0.006*
Visual Complexity x Phonological Cue	-0.25	0.10	-0.45 - -0.05	-2.47	p=0.01*
Visual Complexity x No Phonological Cue	-0.11	0.07	-0.25 - 0.02	-1.60	p=0.11
Random Effects		**Variance**	**SD**	**Correlation**	
Intercepts					
Items		0.32	0.57		
Participants		1.74	1.32		
Slopes (varying over Participants)					
Condition Phonological		0.29	0.54		
Condition Semantic		0.14	0.37	0.77	
Length		0.01	0.12		
Frequency		0.04	0.21		
Imageability		0.03	0.17		
Visual Complexity x Phonological Cue		0.04	0.19		
Visual Complexity x No Phonological Cue		0.01	0.12	1.00	

Key: For each Cue comparison, the reference condition is the one to the right of the 'vs'.

R model equation: Accuracy ~ (1+Condition|Subject) + (1|Word) + Phonological Cue + zLength + zFrequency + zImageability + zNameAgreement + Phonological Cue: zVisualComplexity. Confidence intervals calculated with confint.merMod() function in lme4

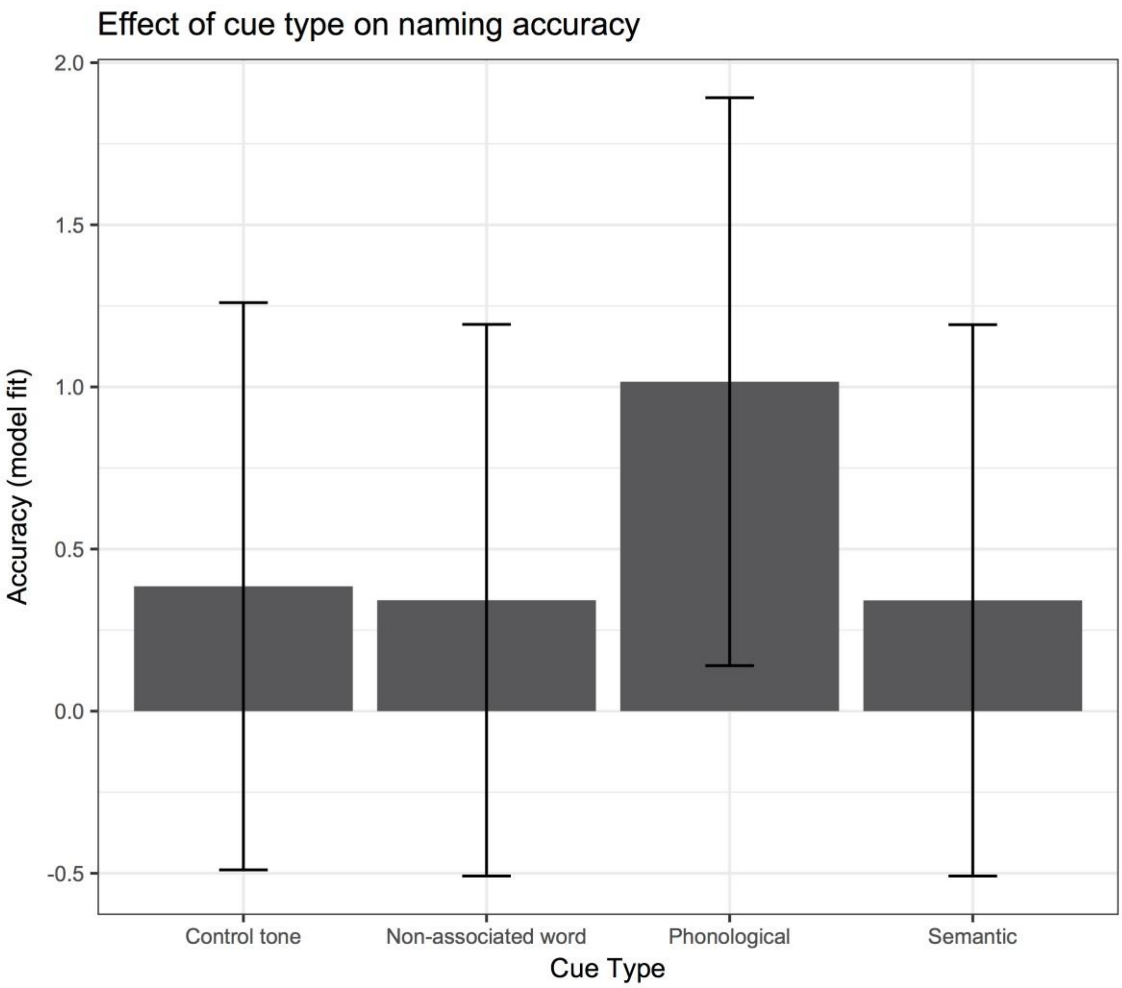

Figure 1. Effect of cue type on naming accuracy. The y axis presents fitted values from the generalized (logistic) linear mixed effect model. Error bars are 95% confidence intervals.

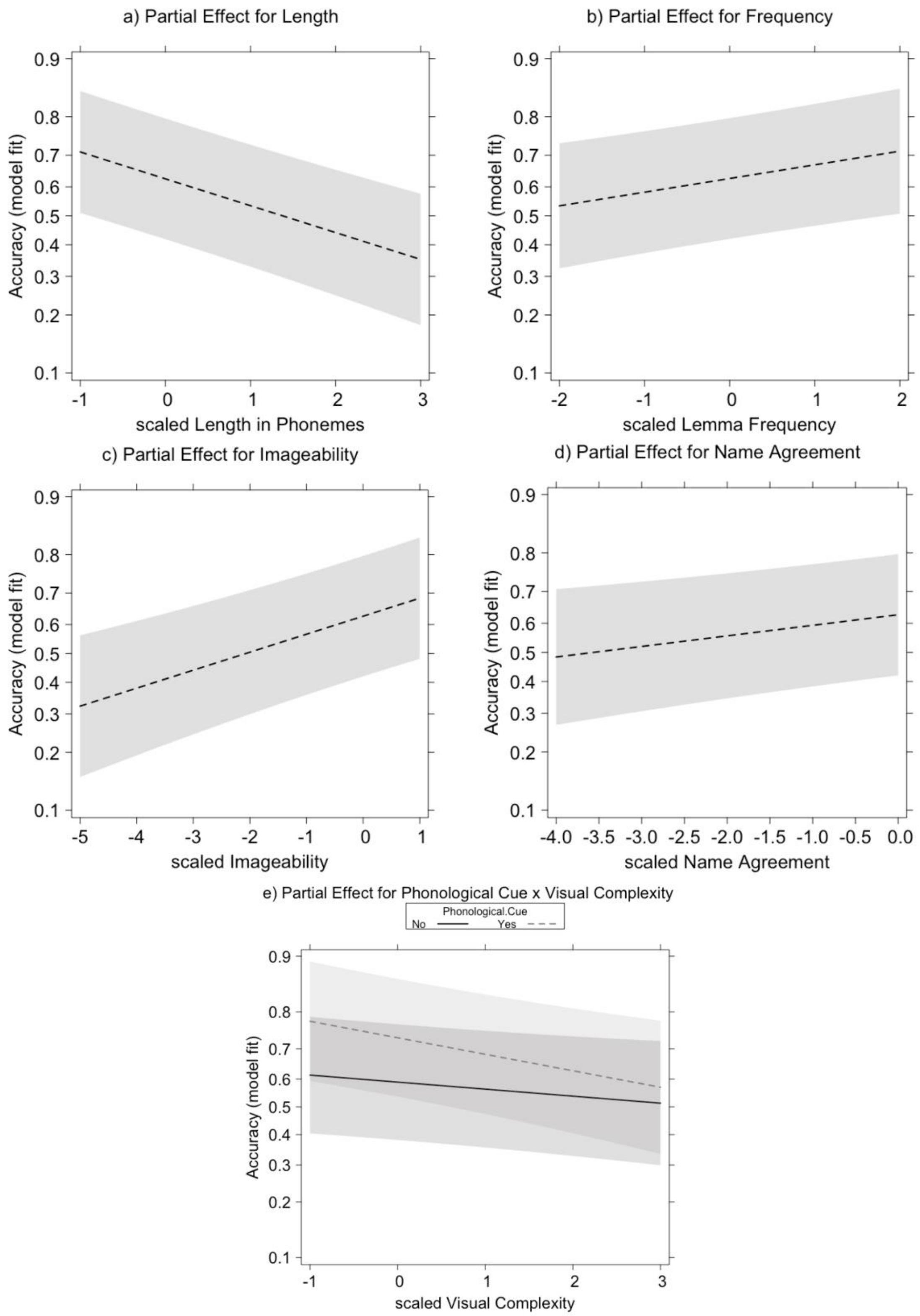

Figure 2. Partial effects plots for the main effects of Length, Frequency, Imageability and Name Agreement, and the interaction of Phonological Cue with Visual Complexity. The y axis presents fitted values from the generalized (logistic) linear mixed effect model. The x axis presents scaled scores for each lexical variable (z scores). Error bars are 95% confidence intervals computed over subject averages.

Table 5. Odds ratios for each participant (taken from model coefficients).

Participant	Intercept	Phonological Cue	Length	Frequency	Imageability	Name Agreement	Visual Complexity x Phonological Cue	Visual Complexity x No Phonological Cue
JV	3.62	1.92	-0.42	0.17	0.03	0.15	-0.2	-0.08
JK	0.86	1.92	-0.4	0.27	0.44	0.15	-0.27	-0.12
DH	0.75	1.92	-0.35	0.15	0.16	0.15	0.03	0.06
WR	2.3	1.92	-0.42	0.1	0.26	0.15	-0.12	-0.03
EW	3.88	1.92	-0.37	-0.16	0.21	0.15	-0.17	-0.06
FF	0.56	1.92	-0.55	0.44	0.18	0.15	-0.31	-0.15
MH	0.36	1.92	-0.24	0.17	0.22	0.15	-0.44	-0.23
AM	0.14	1.92	-0.36	0.38	0.48	0.15	-0.47	-0.25
AW	4.4	1.92	-0.31	0.09	0.2	0.15	-0.38	-0.19
CB	7.97	1.92	-0.45	-0.1	0.43	0.15	-0.14	-0.04
Average		1.92	-0.387	0.151	0.261	0.15	-0.247	-0.109

The odds ratio for the value of the intercept is the odds of a "success" (i.e. the odds of accurate naming) when x = 0 (i.e. 0 values for the other predictors in the model, in this case target properties). The odds ratios for coefficients are the increase (or decrease) in the likelihood / odds of successful naming above the value of the intercept when you add one whole x value (i.e. unit increase in length; presence of phonological cue).

Discussion

In a controlled experiment using within-subject comparison, phonological cues were found to be more effective both in comparison to a control cue and in comparison to semantic cues. This highlights that when task structure, cue provision and response demands were controlled, phonological cues were most effective. The literature shows varying results for which cue (phonological or semantic) is most effective in improving picture naming. Results range from no benefit, to cue to equivalent benefits, to phonological cues to be more effective (e.g., Drew & Thompson, 1999; Lorenz & Ziegler, 2009; Pellet Cheneval et al., 2013; Van Hees et al., 2013). Many therapy studies have demonstrated that phonological cues are not just effective in individuals with phonological processes impairments but also for individuals with semantic impairments. For example, Raymer et al. (1993) showed that four participants with semantic and lexical access impairment improved in naming therapy following phonological cueing. The findings of the current study corroborate recent data showing that phonological cues were more effective in improving picture naming than semantic cues (Lorenz & Ziegler, 2009; Van Hees et al., 2013). Across participants, the presences of a phonological cue made a correct response almost twice as likely.

We also found that semantic cues were not effective when compared against control cue conditions. This is in contrast to a large body of therapy research which had shown semantic cues to be effective ways to improve picture naming, with some earlier studies claiming semantic approaches to be more effective than phonological cues (e.g., Boyle, 2004). This could stem from the type of semantic cue and the manner in which the cues were provided in our experiment. We provided word associate as semantic cues, whilst in semantic therapies cues that are provided could range from yes/no judgement about categorical and attributive information, descriptions, use of an object, and often for a single word a combination of cues are provided (Kiran & Bassetto, 2008). Thus, the multiple sources of semantic information provided for semantic cues in a therapy context is in contrast to a single cue (i.e., word associate) delivered in our experiment. In addition, the word associates were presented auditorily prior to the picture without any opportunity to consider them actively. It is well established that cueing benefits are more effective when PWA engage with them more actively and deeply, be it semantic or phonological cues (e.g., Bose 2013; Hickin et al., 2002; Leonard et al., 2008). Future research testing within participants' comparison of types of semantic cues would shed further light into the differential benefits of various types of cues.

Based on the literature on the effect of lexical properties on picture naming, we anticipated that certain words (short, more frequent, highly imageable, high name agreement) would have greater partial activation and would need less cueing to generate accurate naming. We found positive effects for a number of variables – words that were shorter, more imageable and pictures with higher name agreement were all named more accurately. This is in line with previous literature (e.g., Bose & Schafer, 2017; Kittredge et al., 2008; Middleton & Schwartz, 2010) confirming the important role of both picture and word properties on accurate naming in aphasia. Critically however, the findings support the idea that properties affecting *early* stages of picture identification and lexical retrieval, such as name agreement and visual complexity, are influential in predicting accuracy in picture naming (e.g., Alario et al., 2004; Bose & Schafer, 2017; Ellis & Morrison, 1998; Vitkovitch & Tyrrell, 1995). As images are essential materials for assessments and rehabilitation in aphasia, it is important to determine how image properties affect naming responses. An individual with aphasia may show poor naming performance because the pictures are poor, falsely inflating the measure of their impairment.

We found no interactions between the phonological cue and the majority of target variables, despite finding clear and stable main effects (any effect of frequency was likely unreliable, see Appendix C). Thus, predictions 2a-2c were not supported. This was surprising to us. We predicted that phonological cues would interact with length, since items that are longer are harder to retrieve (as reflected the in the main effect of length) and a phonological cue should directly support retrieval of target phonology. It may be that if phonological cues were provided following picture presentation or after a retrieval failure we may have seen the (in our mind) straightforward mapping between phonological cues and word length.

Visual complexity was the only variable that interacted with phonological cues. Specifically, when a phonological cue was present, items with lower visual complexity were named more accurately than items with high visual complexity. When a phonological cue was not present, there was no effect of visual complexity. Facilitative effects of low visual complexity on picture naming reaction times have been found in healthy individuals (Ellis & Morrison, 1998; Vitkovitch & Tyrrell, 1995), whilst other studies failed to find any effects of visual complexity (e.g., Bates et al., 2003; Bonin et al., 2003). In the healthy literature, longer reaction times to more complex pictures has been attributed to the added detail causing longer picture recognition processes (Ellis & Morrison, 1998). Previous work with PWA has generally either failed to find any effect of visual complexity on output (Nickels & Howard, 1995) or found an effect in the opposite direction, that is, more visually complex items were more likely to be produced correctly (Cuetos, Aguado, Izura & Ellis, 2002). Cuetos et al. (2002) suggest that more detail assisted the recognition of the picture in their patients due to ‘activation of more visual semantic material’ (p. 363) aiding subsequent lexical retrieval. Our findings support this interpretation, since pictures with high visual complexity were named equally accurately whether a phonological cue was present or not. We found that items with less visual information (i.e. lower visual complexity) benefitted from a phonological cue. Thus, pictures of this kind may be hindering naming, as the visual input is sparse and provides less information for easy identification (in line with Cuetos et al, 2002). Phonological cues did not interact with other lexical variables (e.g. length in phonemes, frequency) which could suggest that phonological cues do not affect lexical or output phonology during picture naming – if phonological cues were facilitating the production of phonological information that is partially available, we predicted an interaction of phonological cues with length (i.e. phonological cues aiding the retrieval of longer words) or frequency (i.e. phonological cues aiding the retrieval of items of lower frequency). The finding for visual complexity could be attributed cues being presented prior to the picture. That is, the individual may not be sure that they are seeing an apple, a ball or a balloon, but the provision of the phonological information "*ba*" beforehand rules out apple. Results may have been different if cues were presented after the picture had appeared, or only following a retrieval failure (as in standard assessments of picture naming). However, our data supports literature showing that phonological cues facilitate the mapping from picture concept to word form. In a recent study, Heath et al. (2013) suggested that phonological cues work to effect object recognition in the short term, and strengthen the links from semantics to phonology. During picture naming, we are asking people to recognize a picture and retrieve an appropriate name for that picture. If individuals cannot retrieve a name, provision of a phonological cue helps to constrain the 'search space' that the person is using when they look at the picture (Best et al., 2002; Bose & Buchanan, 2007). In this way, the phonological information is fed back to the early stages of picture recognition, facilitating word retrieval by improving the specificity of conceptual information that is retrieved. This then feeds-forward to lexical and word form retrieval and increases the likelihood of a correct naming response. This interpretation is in line with models of word production that allow for non-linear, cascading of information between different levels of production (i.e. from phonology back to object recognition; see also Griffin & Bock, 1998).

As pointed out by a reviewer, picture recognition errors are typically rare for individuals with aphasia – suggesting that picture recognition processes could be intact and the word retrieval difficulties are driven primarily by problems with accessing lexical information (i.e., lexico-semantic or phonological representations). However, the majority of the literature on picture naming in aphasia has focused on these lexical aspects of naming, rather than consideration of the picture and image properties. Bose & Schafer (2017) showed that pictures with low name agreement resulted in higher error rates for a group of individuals with aphasia. Name agreement arguably involves early processes linked to picture recognition and the flow of information from visual/conceptual semantics to lexical retrieval. Our data shows that phonological cues support naming for pictures which have sparser visual information (i.e. low visual complexity). Whilst it may be the case that picture and object recognition is broadly intact for the majority of individuals with aphasia, if we assume a system that is highly interactive (see above) then lexical retrieval will be more sensitive to variations in the quality of information provided from the image. In other words, an individual can recognize the picture but if the visual information is sparse, noisy or degraded then lexical retrieval will be affected.

In summarize, our data reveals that phonological cues were more effective than semantic cues in improving naming accuracy across individuals. Phonological cues interacted with properties of the picture itself (i.e., visual complexity). The findings challenge the notion of a straightforward mapping from phonological cues to phonological processes. Instead, we see that phonological cues can support naming by feeding back to early picture recognition processes. The data help to explain why patterns of cueing are not consistent in aphasia, i.e., it is not the case that phonological impairments always benefit from phonological cues and semantic impairments form semantic cues. Phonological cues may be more broadly beneficial because they can support the conceptual and semantic information that is retrieved during picture recognition.

Acknowledgments

We are indebted to all our participants for their enthusiasm and time for participation in this research. Data analysis was supported by a British Academy Skill Acquisition Grant (SQ120069) to LM.

References

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.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. J*ournal of Memory and Language, 59*, 390-412.

Baayen, R.H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2).

Baayen, R., Piepenbrock, R., & van Rijn, H. (1993). The CELEX lexical database (CD-ROM). Pennsylvania: Linguistic Data Consortium, University of Pennsylvania, PA.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255-278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014) lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-6. http://CRAN.R-project.org/package=lme4

Baum, S. (1997). Phonological, semantic and mediated priming in aphasia. *Brain & Language, 60,* 347-359.

Best, W. Greenwood, A., Grassly, J., Herbert, R., Hickin, J., & Howard, D. (2013). Aphasia rehabilitation: Does generalization from anomia therapy occur and is it predictable? A case series study. *Cortex, 49*, 2345-2357.

Best, W., Herbert, R., Hickin, J., Osborne, F., & Howard, D. (2002). Phonological and orthographic facilitation of word-retrieval in aphasia: Immediate and delayed effects. *Aphasiology, 16*, 151-168.

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.

Bose, A. (2013). Phonological naming therapy in jargon aphasia with neologisms: effects on picture naming and neologisms. *International Journal of Language and Communication Disorders, 48,* 582-595.

Bose, A., & Buchanan, L. (2007). A cognitive and psycholinguistic investigation of neologisms. *Aphasiology, 21,* 726-738.

Bose, A., & Schafer, G. (2017). Name agreement in aphasia. *Aphasiology*. *31,* 1143-1165.

Boyle, M. (2004). Semantic feature analysis treatment for anomia in two fluent aphasia deficits in aphasia. *Journal of Speech, Language, and Hearing Research, 42*, 972-989.

Conroy, P. J., Snell, C., Sage, K. E., & Lambon Ralph, M. A. (2012). Using phonemic cueing of spontaneous naming to predict item responsiveness to therapy for anomia in aphasia. *Archives of Physical Medicine and Rehabilitation, 93*(1), S53-S60.

Cuetos, F., Aguado, G., & Caramazza, A. (2000). Dissociation of semantic and phonological errors in naming. *Brain and Language*, *75*, 451-460.

Cuetos, F., Aguado, G., Izura, C., & Ellis, A. W. (2002). Aphasic naming in Spanish: predictors and errors. *Brain and Language*, *82*, 344–365.

Davis, A., & Pring, T. (1991). Therapy for word-finding deficits: More on the effects of semantic and phonological approaches to treatment with dysphasic patients. *Neuropsychological Rehabilitation*, *1*, 135-145.

Drew, R. L, & Thompson, C.K. (1999). Model-based semantic treatment for naming deficits in aphasia. *Journal of Speech, Language, and Hearing Research*, *42*, 972-989.

Ellis, A. W., & Morrison, C.M. (1998). Real age-of-acquisition effects in lexical retrieval. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *24,* 515-523.

Fox, J. (2003). Effect Displays in R for Generalised Linear Models. Journal of Statistical Software, 8(15), 1-27. URL http://www.jstatsoft.org/v08/i15/.

Foygel, D., & Dell, G.S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language, 43*, 182-216.

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *Boston Diagnostic Aphasia Examination (BDAE)*. Baltimore, MD : Lippincott Williams & Wilkins.

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, *38*(3), 313-338.

Hagoort, P. (1997). Semantic priming in Broca's aphasics at a short SOA: No support for an automatic access deficit. *Brain & Language, 56*, 287-300.

Heath, S., McMahon, K. L., Nickels, L., Angwin, A., MacDonald, A. D., van Hees, S., Johnson, K., McKinnon, E., & Copland, D. A. (2012) Neural mechanisms underlying the facilitation of naming in aphasia using a semantic task: an fMRI study. *BMC Neuroscience, 13* 98: 98.1-98.19. doi:10.1186/1471-2202-13-98

Heath, S., McMahon, K. L., Nickels, L., Angwin, A., MacDonald, A. D., van Hees, S., Johnson, K., McKinnon, E., Johnson, K., & Copland, D. A. (2013). Facilitation of naming in aphasia with auditory repetition: an investigation of neurocognitive mechanisms. *Neuropsychologia*, *51*, 1534-1548.

Hicken, J., Best, W., Herbert, R., Howard, D., & Osborne, F. (2002). Phonological therapy for word-finding difficulties: A re-evaluation. *Aphasiology, 16*, 981-999.

Howard, D., & Gatehouse, C. (2006). Distinguishing semantic and lexical word retrieval deficits in people with aphasia. *Aphasiology, 20*, 921-950.

Howard, D., & Patterson, K. (1992). *Pyramids and Palm Trees Test*. Bury. St Edmunds: Thames Valley Test.

Howard, D., Hickin, J., Redmond, T., Clark, P., & Best, W. (2006). Re-visiting semantic facilitation of word retrieval for people with aphasia: facilitation yes but semantic no. *Cortex, 42*, 946-962.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: a critical update. *Frontiers in Psychology*, 2 (255), 1-16.

Jefferies, E., & Ralph, M. A. L. (2006). Semantic impairment in stroke aphasia versus semantic dementia: a case-series comparison. *Brain*, *129*, 2132-2147.

Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic assessments of language processing in aphasia.* London: Psychology Press.

Kiran, S., & Bassetto, G. (2008). Evaluating the effectiveness of semantic based treatment for naming deficits in aphasia: what works? *Seminars in Speech and Language, 29*, 71-82.

Kittredge, A., Dell, G., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology, 25*, 463– 492.

Kuznetsova, A., Brockhoff, P.B. & Bojesen, R.H. (2014). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2.0-6. http://CRAN.R-project.org/package=lme4

Laine, M., & Martin, N. (2006). *Anomia: Theoretical and clinical aspects*. Hove, UK: Psychology Press.

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1-75.

Li, E.C., & Williams, S.E. (1991). An investigation of naming errors following semantic and phonemic cueing. *Neuropsychologia, 29*, 1083-1093.

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, *22*, 1-35.

Lorenz, A., Ziegler, W. (2009). Semantic vs. Word Form Specific Techniques in Anomia Treatment: A Multiple Single Case Study. *Journal of Neurolinguistics*, 22, 515-537

Lorenz, A., & Nickels, L. (2007). Orthographic cueing in anomic aphasia: How does it work? *Aphasiology, 21*, 670-686.

Lupyan, G. (2008). The conceptual grouping effect: Categories matter (and named categories matter more). *Cognition*, *108*, 566-577.

Middleton, E. L., & Schwartz, M. F. (2010). Density pervades: an analysis of phonological neighbourhood density effects in aphasic speakers with different types of naming impairment. *Cognitive Neuropsychology*, *27*, 401–27.

Morris, S. B., & DeShon, R. P. (1997). Correcting effect sizes computed from factor analysis of variance for use in meta-analysis.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://www.usf.edu/FreeAssociation/.

Nickels, L. (2002). Theoretical and methodological issues in the cognitive neuropsychology of spoken word production. *Aphasiology, 16*, 3-19.

Nickels, L., & Howard, D. (1995). Phonological errors in aphasic naming: Comprehension, monitoring and lexicality. *Cortex*, *31*, 209 - 237.

Nickels, L., & Best, W. (1996). Therapy for naming deficits (part I): Principles, puzzles and progress. *Aphasiology*, *10*, 21–47.

Pellet Cheneval, P., Bonnans, C., & Laganaro, M. (2017). Does facilitation by phonological cuing in picture naming depend on the modality of the cue?. Aphasiology, DOI: 10.1080/02687038.2017.1328102

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107*, 460-499.

Raymer, A.M., Thompson, C.K., Jacobs, B., & Le Grand, H.R. (1993). Phonological treatment of naming deficits in aphasia: model-based generalization analysis. *Aphasiology, 7*, 27-53.

Reitan, R. M., & Wolfson, D. (1993). *The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation* (2nd ed.). S. Tuscon, AZ: Neuropsychology Press.

Roach, A., Schwartz, M. F., Martin, N., Grewal, R.S., & Brecher, A. (1996). The Philadelphia Naming Test: Scoring and Rationale. *Aphasiology*, *24*, 121-133.

Schwartz, M. F. (2013). Theoretical analysis of word production deficits in adult aphasia. *Philosophical Transactions of the Royal Society B*, 369: 20120390.

Stimely, M.A., & Noll, J. D. (1991). The effects of semantic and phonemic prestimulation cues on picture naming in aphasia. *Brain and Language*, *41*, 496-509.

Székely, A., & Bates, E. (2000). Objective visual complexity as a variable in studies of picture naming. *Center for Research in Language Newsletter*, *12*, 3-33.

Van Hees, S., Angwin, A., McMahon, K., & Copland, D. (2013). A comparison of semantic feature analysis and phonological components analysis for the treatment of naming impairments in aphasia. *Neuropsychological Rehabilitation, 23*, 102-132.

Vitevitch, M. S., & Sommers, M. S. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production in younger and older adults. *Memory & Cognition, 31*, 491–504.

Vitkovitch, M., & Tyrrell, L. (1995). Sources of disagreement in object naming. *The Quarterly Journal of Experimental Psychology*, *48*, 822-848.

Wickham, R. (2009). ggplot2: elegant graphics for data analysis. Springer New York, 2009.

Wisenburn, B., & Mahoney, K. (2009). A meta-analysis of word-finding treatments for aphasia. *Aphasiology*, *23*, 1338-1352.

APPENDIX A

This appendix provides the Philadelphia Naming Test (Roach et al., 1996) stimulus list with the lexical properties used in this study namely: length (in phonemes), frequency (lemma, per million), imageability, name agreement and visual complexity. We also provide the words used for semantically related cue and non-associated control cue.

Item #	Target Word	Length (Phonemes)	Lemma (base 10) Log Frequency per million (CELEX)	Imageability	Name agreement	Visual complexity	Semantically related cue	Semantically non-associated control cue
1	candle	6	1.2041	6.54	1	129	wick	chop
2	ghost	4	1.4914	5.29	1	211	scary	insult
3	dinosaur	7	0.699	6.11	1	205	extinct	pretend
4	tree	3	2.281	6.58	1	303	sap	sill
5	pen	3	1.415	6.37	1	98	ink	hedge
6	scissors	6	0.6021	6.58	1	138	cut	late
7	cane	3	1	5.32	1	76	walk	write
8	comb	3	0.699	6.11	1	330	hair	rose
9	thermometer	10	0.7782	5.82	0.9	139	temperature	dictionary
10	well	3	0.699	4.42	1	246	bucket	letter
11	grapes	4	1	6.55	1	182	vine	chick
12	strawberries	8	0.7782	6.66	0.9	199	shortcake	illness
13	bread	4	1.8692	6.24	1	188	loaf	hut
14	football	6	1.5185	6.58	1	174	touchdown	permit
15	pig	3	1.6335	6.45	1	128	ham	dent
16	apple	4	1.4771	6.55	1	120	orchard	carpet
17	hand	4	2.8597	6.5	1	140	glove	deer
18	towel	4	1.3424	6.37	0.8	160	bath	rice
19	lion	4	1.3979	6.63	0.95	255	roar	toll
20	glass	4	2.1614	5.84	1	173	shatter	publish
21	fork	4	1.1761	6.45	1	118	spoon	calf
22	plant	5	2.0828	6.03	0.95	215	seed	match
23	garage	5	1.3979	5.82	1	386	storage	partner
24	can	3	0.9542	4.71	1	159	opener	stereo
25	table	5	2.3711	6.42	1	173	chair	chin
26	waterfall	8	0.9031	6.42	1	365	pool	plane
27	king	3	1.9956	5.73	1	334	throne	rash
28	boot	3	1.5911	6.18	1	138	kick	rash
29	foot	3	2.5132	6.37	1	90	toe	cat
30	chair	3	2.1335	6.47	1	191	table	road
31	banana	6	0.9031	6.84	1	175	monkey	package
32	ring	3	1.6902	5.95	1	113	diamond	lotion
33	dice	3	0.301	6.05	1	260	gamble	refresh
34	calendar	8	0.9031	5.58	1	277	date	buy
35	knife	3	1.6435	6.39	1	112	fork	joke
36	vest	4	0.8451	6.03	1	185	sweater	apple
37	turkey	5	0.699	6.26	1	304	thanksgiving	fascinate
38	rake	3	0.301	5.84	0.95	148	leaves	cub
39	balloon	5	0.7782	6.5	1	87	helium	inferior
40	duck	3	1.1461	6.34	1	265	quack	chess
41	fireplace	7	0.9542	6.32	1	185	chimney	gamble
42	pineapple	7	0.4771	6.58	1	297	fruit	tie

43	fan	3	1.2304	5.68	1	288	air	eye
44	window	5	2.301	6.16	1	299	pane	owl
45	lamp	4	1.5441	6.05	1	100	desk	soil
46	drum	4	1.2041	6.03	1	345	beat	cool
47	skull	4	1.3222	6.29	0.75	216	brain	tool
48	bridge	4	1.8195	6.18	1	526	river	women
49	eskimo	6	0.301	5.82	0.9	227	arctic	filter
50	dog	3	2.0607	6.76	1	160	pet	shoe
51	iron	4	1.8513	5.76	1	238	crease	peep
52	cheerleaders	8	0	6.03	0.95	324	football	golden
53	snake	4	1.3617	6.45	1	263	bite	tide
54	ambulance	9	0.9542	6.53	0.9	373	emergency	territory
55	carrot	5	0.9031	6.5	1	149	vegetable	graduation
56	sailor	5	1.0792	5.84	0.95	138	ship	core
57	book	3	2.6375	6.47	1	230	library	orchestra
58	bus	3	1.8976	6.32	0.95	291	tour	shell
59	map	3	1.6021	6.45	1	494	directions	terminate
60	squirrel	7	0.7782	6.39	1	236	acorn	mustard
61	microscope	9	0.9031	5.97	0.84	178	biology	strategy
62	bowl	3	1.5185	6.27	0.85	157	dish	fox
63	van	3	1.7634	6.18	0.95	296	move	heart
64	helicopter	10	1.2041	6.47	1	231	flying	crying
65	bottle	5	2.0645	6.39	0.95	100	wine	team
66	scarf	5	1.0792	6.26	1	209	wool	chop
67	ball	3	2.0453	6.37	1	226	racquet	eating
68	frog	4	0.9542	6.47	1	179	croak	gash
69	cow	2	1.6021	6.5	1	193	moo	oar
70	beard	4	1.3979	6.11	1	399	whiskers	finance
71	glove	4	1.2788	6.32	1	146	hand	old
72	owl	2	0.8451	6.42	1	235	hoot	soak
73	pipe	3	1.4914	5.63	1	139	tobacco	stereo
74	scale	4	1.9138	4.24	0.94	239	measure	friendly
75	tent	4	1.6435	6.58	1	168	camping	chicken
76	flashlight	7	0.699	5.89	1	173	battery	discover
77	camel	5	1.3979	6.19	1	184	desert	blanket
78	goat	3	1.4472	6.24	1	202	mountain	table
79	fish	3	2.2122	6.26	1	212	catch	guide
80	cannon	5	0.7782	6.05	0.95	183	blast	shoe
81	shoe	2	1.8976	6.47	1	150	sock	owl
82	sandwich	7	1	6.34	0.9	238	bread	aim
83	spider	6	0.8451	6.54	1	290	web	cough
84	belt	4	1.4314	6.24	1	211	pants	nasal
85	toilet	5	1.4472	6.5	1	182	flush	calf
86	wagon	5	1.0414	5.43	1	279	pull	add
87	ruler	5	1.2553	6.03	1	88	measure	married
88	tractor	7	1.0414	6.29	0.95	206	farmer	stomach
89	queen	4	1.7243	6.16	1	354	king	edge
90	train	4	1.9085	6.42	1	368	railroad	blanket
91	church	4	2.2625	6.45	1	340	steeple	monkey
92	anchor	5	0.7782	6.18	1	221	boat	cool
93	whistle	5	0.9542	5.58	0.9	158	blow	cape
94	corn	4	1.3802	5.26	1	235	husk	chess
95	pyramid	7	0.8451	6.43	0.95	298	Egypt	tulip
96	typewriter	8	1.0414	6	1	289	secretary	automobile
97	rope	3	1.6232	6.08	1	303	knot	dent
98	basket	6	1.3802	5.87	1	300	picnic	predict
99	letter	5	2.3139	5.66	1	299	stamp	tap

100	nose	3	1.9085	6.21	1	59	snot	corn
101	chimney	5	1	6.24	0.95	331	fireplace	dedicate
102	horse	4	2.1206	6.5	1	232	saddle	spoken
103	key	2	1.9345	6.08	1	160	lock	tie
104	fireman	6	0.6021	6.13	1	237	rescue	harvest
105	cross	4	1.3424	5.47	1	100	holy	bus
106	crutches	4	0.6021	6.13	1	210	walk	sleep
107	bone	3	1.8388	5.92	1	164	skeleton	dominate
108	cat	3	1.8261	6.66	1	171	meow	hose
109	kitchen	5	2.0453	6.08	1	348	apron	strategy
110	dragon	6	0.9542	6.13	1	394	puff	tart
111	saddle	5	1	5.79	0.9	201	horse	fine
112	pie	2	1.2304	6.13	1	156	crust	ache
113	snail	4	0.6021	6.47	1	154	shell	van
114	pirate	5	0.699	6.45	1	464	capture	rental
115	clock	4	1.5911	6.47	1	283	time	head
116	pumpkin	7	0.301	6.45	1	200	Halloween	determine
117	sock	3	1.2553	6.24	1	110	shoe	tin
118	closet	6	1.0414	5.16	1	247	hanger	bullet
119	hair	3	2.2989	6.13	1	458	brush	tape
120	baby	4	2.4116	6.24	1	177	crib	hook
121	bat	3	1.1461	5.84	1	180	vampire	purely
122	leaf	3	1.9085	6.18	1	144	maple	monster
123	slippers	6	0.9542	6.05	0.75	123	feet	read
124	mountain	6	1.9243	6.37	0.95	270	climber	garbage
125	sun	3	2.1818	6.63	0.9	218	rays	shy
126	moustache	6	0	6.16	1	216	beard	hen
127	ear	2	1.9445	6.32	1	161	lobe	harp
128	door	3	2.5866	6.24	1	266	knob	peel
129	house	3	2.7825	6.45	0.95	179	brick	dish
130	nail	3	1.3979	6.05	1	81	hammer	sofa
131	binoculars	11	0.699	6.24	0.9	188	birds	wig
132	celery	6	0.4771	5.68	0.75	233	stalk	ring
133	vase	3	0.8451	5.86	1	203	flower	motor
134	pencil	6	1.2788	6.34	1	136	eraser	unspeakable
135	elephant	7	1.3802	6.63	1	233	tusk	fold
136	hose	3	0.6021	5.84	1	229	garden	vision
137	bench	4	1.415	5.92	1	290	park	text
138	zebra	5	0.301	6.59	1	286	stripe	eel
139	man	3	3.2119	5.82	0.95	189	lady	ship
140	seal	3	1.1461	5.82	0.9	171	walrus	apple
141	wig	3	1.1139	5.58	1	217	hair	rock
142	necklace	6	0.6021	6.32	1	136	pearl	wing
143	desk	4	1.959	5.95	0.95	238	office	water
144	bell	3	1.6232	6.08	1	151	chime	hair
145	star	4	2.0043	5.95	1	114	astronomy	graduation
146	hammer	5	1.0414	6.32	1	133	nail	ape
147	pillow	4	1.2788	6.24	1	115	sheets	spoon
148	spoon	4	1.1761	6.34	1	142	fork	wool
149	zipper	5	0.301	5.84	0.95	95	button	daisy
150	top	3	2.2355	3.95	0.95	120	spin	tame
151	flower	6	1.9685	6.29	1	214	petals	bullet
152	kite	3	0.699	6.11	1	158	flying	sleeping
153	suit	3	1.716	5.92	1	201	tie	hen
154	cake	3	1.5315	6.55	1	185	icing	invest
155	hat	3	1.8325	6.39	1	128	coat	moon
156	crown	4	1.3802	6.08	1	318	throne	eel

157	piano	5	1.4314	6.42	0.95	305	music	single
158	stethoscope	9	0	6.13	1	172	doctor	window
159	bride	4	1.0792	6.24	0.94	396	groom	tin
160	butterfly	8	1	6.35	1	421	cocoon	scored
161	heart	4	2.2148	6.13	1	101	valentine	jockey
162	skis	3	0.9031	6	1	216	slope	pie
163	clown	4	0.6021	6.11	1	398	circus	filter
164	volcano	7	0.7782	6.49	1	366	erupt	boxer
165	pear	3	0.7782	6.47	1	84	fruit	bus
166	octopus	7	0.301	6.39	1	392	tentacles	recording
167	saw	2	0	5.79	1	121	chain	bird
168	camera	6	1.5563	6.32	1	231	tourist	harvest
169	bed	3	2.4298	6.58	1	192	sleep	fort
170	harp	4	0.4771	6.08	0.94	223	music	single
171	broom	4	0.9031	6.13	0.9	181	sweep	thaw
172	nurse	4	1.6902	6.11	1	189	doctor	colour
173	eye	1	2.7185	6.42	1	221	sight	thin
174	cowboy	4	0.7782	5.89	0.95	227	horse	book
175	monkey	5	1.2553	6.55	1	188	banana	rejection

APPENDIX B

This Appendix provides details of additional models we ran following the recommendation of the reviewers with random intercepts and correlated random slopes (reported below). These models do not change the key findings, with the only difference being that the main effect for Frequency was no longer significant. Note the perfect correlations between intercepts and slopes, indicating that the addition of slopes are not explaining additional variance in the model. No intercept was included for imageability as models including both intercepts and slopes for imageability did not converge.

Model refit with data points excluded that had residuals >2.5 standard deviations

Fixed Effects	**Estimate**	**Standard Error**	**Wald 95% CI**	**Z**	***p***
Intercept	0.30	0.55	-0.78 - 1.37	0.54	0.59
Phonological Cue vs No Phonological Cue	0.82	0.10	0.63 - 1.01	8.56	p<0.001*
Length	-0.50	0.09	-0.68 - -0.32	-5.38	p<0.001*
Frequency	0.23	0.11	0.03 - 0.44	2.22	p=0.03*
Imageability	0.36	0.11	0.15 - 0.56	3.37	p<0.005*
Name Agreement	0.18	0.07	0.05 - 0.31	2.68	p=0.007*
Visual Complexity x Phonological Cue	-0.33	0.12	-0.57 - -0.09	-2.74	p=0.006*
Visual Complexity x No Phonological Cue	-0.16	0.09	-0.35 - 0.02	-1.71	p=0.09
Random Effects		**Variance**	**SD**	**Correlation**	
Intercepts					
Items		0.52	0.72		
Participants		2.75	1.66		
Slopes (varying over Participants)					
Condition Phonological		0.67	0.82		
Condition Semantic		0.14	0.38	0.99	
Length		0.03	0.16		
Frequency		0.06	0.24		
Imageability		0.06	0.25		
Visual Complexity x Phonological Cue		0.06	0.24	1.00	
Visual Complexity x NoPhonological Cue		0.04	0.20		

"R model equation: ACC ~ (1 | Subject) +(1 | Word) + (0 + Condition | Subject) + (0 + zLengthPh | Subject) + (0 + zFreq | Subject) + (0 + zImageability | Subject) + (0 + PhCue:zVisComp | Subject) + PhCue + zLengthPh + zFreq + zImageability + zNameAgr + PhCue:zVisComp"
Confidence intervals calculated with confint.merMod() function in lme4

Following Baayen & Milin (2010), we looked at the number of data points for which standardised residuals were greater than 2.5 or less than -2.5; there were 38 such data points. The advice in Baayen & Milin is to refit a model after removing data points for there are large residuals (>2.5). If the model parameters and coefficients are the similar following refitting, we can argue that findings in the original model were not overly influenced by outliers or data points leveraging model estimates.

APPENDIX C

Maximal model (random intercepts and correlated random slopes)

Fixed Effects	**Estimate**	**Standard Error**	**95% CI**	**Z**	***p***
Intercept	0.30	0.43	-0.88 - 1.51	0.71	0.48
Phonological Cue vs No Phonological Cue	0.65	0.09	0.42 - 0.88	7.50	p<0.001*
Length	-0.39	0.08	-0.59 - -0.19	-5.21	p<0.001*
Frequency	0.15	0.09	-0.09 - 0.40	1.63	p=0.10
Imageability	0.26	0.08	0.05 - 0.47	3.42	p<0.001*
Name Agreement	0.15	0.06	0.00 - 0.30	2.74	p=0.006*
Visual Complexity x Phonological Cue	-0.25	0.10	-0.53 - 0.03	-2.42	p=0.02*
Visual Complexity x No Phonological Cue	-0.11	0.07	-0.30 - 0.07	-1.60	p=0.11
Random Effects		**Variance**	**SD**	**Correlation**	
Intercepts					
Items		0.32	0.56		
Participant Intercepts & Slopes					
Participant x Condition (intercept)		0.27	0.52		
Condition (slope)		0.12	0.35	-1.00	
Participant x Length (intercept)		0.29	0.54		
Length (slope)		0.16	0.13	-1.00	
Participant x Frequency (intercept)		1.45	1.20		
Frequency (slope)		0.05	0.22	-1.00	
Imageability (slope)		0.03	0.17		
Participant x Name Agreement (intercept)		0.00	0.00		
Name Agreement (slope)		0.00	0.00	1.00	
Participant x Visual Complexity x Phonological Cue (intercept)		0.01	0.10		
Visual Complexity x Phonological Cue (slope)		0.04	0.20	1.00	
Visual Complexity x NoPhonological Cue (slope)		0.01	0.12	1.00	1.00

R model equation: Accuracy ~ (1 + Condition | Subject) + (1 | Word) + (1 + zLengthPh | Subject) + (1 + zFreq | Subject) + (0 + zImageability | Subject) + (1 + zNameAgr | Subject) + (1 + PhCue:zVisComp | Subject) + PhCue + zLengthPh + zFreq + zImageability + zNameAgr + PhCue:zVisComp
Confidence intervals calculated with confint.merMod() function in lme4