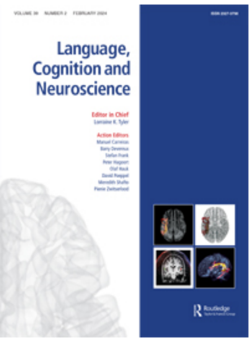

Language, Cognition and Neuroscience

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/plcp21

Cascaded processing develops by five years of age: evidence from adult and child picture naming

Margaret Kandel & Jesse Snedeker

To cite this article: Margaret Kandel & Jesse Snedeker (2024) Cascaded processing develops by five years of age: evidence from adult and child picture naming, Language, Cognition and Neuroscience, 39:2, 167-202, DOI: 10.1080/23273798.2023.2258536

To link to this article: https://doi.org/10.1080/23273798.2023.2258536

LANGUAGE, COGNITION AND NEUROSCIENCE
2024, VOL. 39, NO. 2, 167–202
https://doi.org/10.1080/23273798.2023.2258536

Routledge
Taylor & Francis Group

REGULAR ARTICLE

Cascaded processing develops by five years of age: evidence from adult and child picture naming

Margaret Kandel and Jesse Snedeker

Department of Psychology, Harvard University, Cambridge, MA, USA

ABSTRACT
Although there is compelling evidence for cascading activation in adult lexical planning, there is little research on how and when cascaded processing develops. We use a picture naming task to compare word planning in adults and five-year-old children. We manipulated image codability (name agreement) and name frequency, factors that affect lexical selection and phonological encoding, respectively. These factors had qualitatively similar influences on naming response time in both populations, suggesting similar underlying planning processes. Critically, we found an under-additive interaction between codability and frequency such that the frequency effect was attenuated when name agreement was low. This interaction generalises across experiments and languages and can be simulated in a planning architecture in which phonological forms become activated before lexical selection is complete. These results provide evidence for cascaded processing at an earlier age than previous studies, suggesting that informational cascades are a fundamental property of the production architecture.

ARTICLE HISTORY
Received 31 January 2023
Accepted 28 August 2023

KEYWORDS
Language production; cascaded processing; word planning; name agreement; codability; frequency; language acquisition

Introduction

Psycholinguistic theories break the process of producing a word into several steps (e.g. Butterworth, 1989; Dell, 1986; Friedmann et al., 2013; Garrett, 1980; Levelt, 1989; Levelt, 2001; Schriefers & Vigliocco, 2015; inter alia). Contemporary models have three major stages prior to articulation planning: (i) conceptual processing (determining the lexical concept to be conveyed; e.g. Levelt et al., 1999), (ii) lexical selection (identifying the lexical representation from the mental lexicon that will best convey the lexical concept), and (iii) phonological encoding (accessing the phonological form of the word to be uttered). A central question in language production research is how these processing levels interact and how information is passed between them.

There is compelling evidence that adult speakers carry out these processes in a cascaded fashion, with one process beginning before the earlier one is complete. Specifically, speakers can begin accessing the phonological forms of the lexical items under consideration even before a particular lexical item has been selected (e.g. Costa et al., 2000; Cutting & Ferreira, 1999; Jescheniak & Schriefers, 1998; Morsella & Miozzo, 2002; Peterson & Savoy, 1998; Rapp & Goldrick, 2000; Starreveld & La Heij, 1995). This process is frequently modelled as activation spreading between networks of interconnected nodes (e.g. Bloem & La Heij, 2003; Caramazza, 1997; Collins & Loftus, 1975; Dell, 1986; Dell et al., 1997; Levelt et al., 1999; Roelofs, 1992). Cascaded processing has a potential functional benefit: speakers can get a head start on later levels of processing while finishing up the earlier ones, allowing for rapid and fluent production. In addition, cascaded processing allows representations at lower levels to become active while representations are still under consideration at a higher level, opening the door for processing at lower levels to influence selection at the higher level through feedback loops (e.g. Dell, 1986; Dell et al., 1997; Dell & O'Seaghdha, 1991, 1992; Harley, 1993).

Given the centrality of cascaded processing to our theories of production, understanding the development of this ability is critical. Is cascading activation a fundamental property of the language production system, a consequence of its architecture that is present from early in development? Or does it emerge gradually with experience, appearing only after processing speed (or efficiency) approaches adult-like levels, thus freeing up resources for more processes to occur simultaneously? Curiously, despite the vast literature investigating activation flow in adult word planning, there is

CONTACT Margaret Kandel mkandel@g.harvard.edu Department of Psychology, Harvard University, 33 Kirkland St, Cambridge, MA 02138, USA

This article was originally published with errors, which have now been corrected in the online version. Please see Correction (https://doi.org/10.1080/23273798.2025.2444783)

little work that explores how and when this ability develops (but see Jescheniak et al., 2006 discussed below). The present paper addresses these questions by comparing single word production in adults and five-year-old children using a picture naming paradigm.

Evidence for cascading activation in adult language production

Evidence for cascading activation in adult language production is primarily derived from two major sources: speech error analyses and reaction time studies.

Cascading activation models accurately capture the distributions of speech error types in adults with and without impaired lexical access in a variety of tasks including picture naming, word–picture mapping, and word repetition (e.g. Dell et al., 1997; Foygel & Dell, 2000; Rapp & Goldrick, 2000; Schwartz et al., 2006). Critically, this theory correctly predicts the frequency of *mixed errors*, substitution errors that are both semantically and phonologically related to the intended word (e.g. saying *rat* in place of *cat*). In both spontaneous and experimentally-elicited speech, adults produce mixed errors at a higher rate than would be expected given the base rates of purely semantic and purely phonological errors (e.g. Blanken, 1998; Butterworth, 1981; Dell & Reich, 1981; Harley, 1984; Martin et al., 1989; Martin et al., 1996; Rapp & Goldrick, 2000; Stemberger, 1983). The high likelihood of mixed errors arises as a natural consequence of cascading activation (e.g. Rapp & Goldrick, 2000) but is not predicted by discrete, serial word planning models that assume that a word's phonological form is only activated once its lexical representation has been selected for articulation (e.g. Butterworth, 1992; Garrett, 1980; Levelt et al., 1999; Roelofs, 1992; Schriefers et al., 1990).[1]

Additional evidence for cascading activation comes from reaction time studies that explore when lexical and phonological representations become active during word planning. These studies often use the picture-word interference (PWI) paradigm, in which participants must name a picture accompanied by a distractor word (presented either visually or auditorily). Distractors that are semantically related to the picture referent generally result in slower naming times (e.g. Damian & Bowers, 2003; Damian & Martin, 1999; La Heij et al., 1990; Roelofs, 1992; Schriefers et al., 1990; Vigliocco et al., 2004; but cf. Mahon et al., 2007), an effect which is believed to reflect increased difficulty choosing the intended word during lexical selection due to competition from the semantic distractor (Levelt et al., 1999; Roelofs, 1992; but cf. Mahon et al., 2007 for an alternative proposal). When a distractor is phonologically related to the picture name, on the other hand, naming RTs tend to be shorter (e.g. Levelt et al., 1991; Meyer & Schriefers, 1991; Schriefers et al., 1990), which is thought to reflect activation of the phonemes in the target name shared by the distractor word (Damian & Martin, 1999; Dell & O'Seaghdha, 1992; Roelofs et al., 1996). There is evidence that phonological facilitation can occur within the same early time windows as semantic interference (e.g. Cutting & Ferreira, 1999; Damian & Martin, 1999; Jescheniak & Schriefers, 2001; Starreveld, 2000), suggesting that both lexical and phonological representations can be active at the same time, as predicted by cascading models. In addition, when distractors are both semantically and phonologically related to the target, there is an interaction between the two effects, a pattern which suggests that lexical selection and phonological encoding are not serial and independent (Damian & Martin, 1999; Starreveld & La Heij, 1995). Not all studies have found evidence of early phonological effects, however, (e.g. Schriefers et al., 1990), and these effects do not provide fully conclusive evidence of cascaded processing, as phonological effects in PWI paradigms are open to a range of explanations, including some in which the effect does not arise directly from production processes (Starreveld, 2000). Furthermore, it is possible to account for the interactions observed in PWI tasks between semantic and phonological effects within a serial framework of word planning (e.g. Roelofs et al., 1996).

More direct evidence for cascading activation in adult production is derived from RT studies that probe the phonological activation of un-uttered words, typically unselected alternative lexical candidates or words that are semantically related to the produced word. Evidence for the activation of words related to the produced word (semantically-mediated phonological activation) has been found in a number of tasks including priming paradigms (e.g. Peterson & Savoy, 1998), PWI (e.g. Abdel Rahman & Melinger, 2008; Jescheniak et al., 2005; Jescheniak et al., 2006; Jescheniak & Schriefers, 1998; Melinger & Abdel Rahman, 2013; Zhang et al., 2018; inter alia), and EEG paradigms (Jescheniak et al., 2003). Evidence has also been found for the activation of words related to a homophone of the produced word (Cutting & Ferreira, 1999). Additional evidence for phonological activation of un-uttered words comes from bilingual picture naming (RTs are faster when the names in both languages are phonologically similar, suggesting both phonological forms are activated; Costa et al., 2000) and from picture–picture interference paradigms in which participants have to name one of two superimposed images (RTs are faster when the name of the un-named image is phonologically related to the produced name; e.g. Humphreys et al., 2010; Kuipers & La Heij, 2009; Mädebach et al., 2011; Meyer & Damian, 2007; Morsella & Miozzo, 2002; Navarette et al., 2017; Navarrete & Costa, 2009; Roelofs, 2008).

The present study uses a naming paradigm to explore an additional, unstudied, prediction of the cascading architecture: that phonological activation can begin while lexical selection is still underway, thereby resulting in interactions between the variables that influence each of the two processes (more below). The picture naming task is considerably simpler than the priming and interference tasks discussed above, thus we can investigate this prediction not only in mature speakers but also in young children, for whom the dynamics of lexical access have not been as extensively studied.

The limited evidence for cascading activation in child language production

Speech error and aphasia research suggests that the developing language production system is similarly organised to that of adults, with distinct stages for lexical selection and phonological processing (Friedmann et al., 2013). There is compelling (although limited) evidence for cascading activation in word planning in children over the age of seven. Jescheniak et al. (2006) observed evidence of semantically-mediated phonological activation in the picture naming behaviour of second graders (aged 7;3–8;6). In an auditory PWI task, Jescheniak et al. (2006) found that second graders were slower to produce target picture names (e.g. *Mantel* [coat]) in the presence of a distractor (e.g. *Honig* [honey]) that had the same phonological onset as a word that was semantically related to the target (e.g. *Hose* [trousers]) but was not phonologically or semantically related to the target itself or semantically related to the target-related word. Such an interference effect would be expected if activation cascades to the phonological forms of semantic associates of the target word (Jescheniak et al., 2005; Jescheniak & Schriefers, 1998). Jescheniak et al. (2006) did not find a similar effect in fourth graders (aged 9;4–10;8) or adults, leading them to propose that cascading activation is present across development but is easier to detect when the lexical planning process is more stretched out in time, as it is for the seven and eight-year-old children.

Other studies with school-aged children are suggestive of cascaded processing, even if they are open to alternative explanations. For example, children eight to eleven years old, like adults, show phonological facilitation effects in the same early time windows as semantic interference effects (e.g. Sieger-Gardner & Schwartz, 2008; Sylvia, 2017). In addition, like adults, school-aged children (eight to eleven years) are influenced by the phonological forms of context picture names in picture–picture interference paradigms (Sylvia, 2017), though Sylvia (2017) observed an interference effect from the distractor image rather than the facilitation effect that is commonly observed in adults (e.g. Morsella & Miozzo, 2002). Finally, Poarch and van Hell (2012) found that multilingual children aged five to eight years old (M age = 7.28 years, SD = 0.76) demonstrate a bidirectional cognate phonological facilitation effect between German and English, with faster RTs when the names in both languages are phonologically similar, which *could* indicate that both phonological forms become activated (or it could reflect a phonological frequency confound; see Costa et al., 2000 for discussion).

In contrast, the evidence of cascading activation for children under the age of seven is limited and weak. Three lines of research are potentially relevant. First, children between five and seven years of age show early phonological facilitation effects in an auditory version of the PWI paradigm (Jerger et al., 2002). These effects, however, are open to the same alternative explanations as the parallel effects in adults (e.g. Starreveld, 2000). Second, like adults, young children (1–5 years) produce mixed lexical substitution errors (Jaeger, 2005). Unfortunately, these analyses do not assess whether the frequency of substitution errors both semantically and phonologically related to the target is greater than what would be expected if phonological and semantic errors are independent (as they would be in a model with no informational cascade). Third, adult speech error models that include cascading activation (e.g. Foygel & Dell, 2000) can simulate error distributions in children five to eleven years old with quantitative shifts in model parameters across ages (Budd et al., 2011), suggesting that broadly similar processes are at work in adults and children. Evidence consistent with one model, however, does not rule out others. It is unclear whether the cascading model provides a better fit of the child error distribution than models with no informational cascade.

In sum, while a cascading architecture is likely present by around seven years of age, the evidence for cascading activation in word planning before this age is weaker and open to multiple interpretations. The present study explores whether cascading activation is present for five-year-olds, children who are proficient speakers of their native language but have had little formal schooling and are largely pre-literate.

Reasons why activation flow may differ in child language production

The cognitive and linguistic abilities of children five years of age and younger differ from older children and adults in several ways that could have implications for the development of cascaded processing.

Some of these differences might lead us to expect cascading activation to be limited, weaker, or more difficult to detect in children this young. For example, although five-year-olds are proficient speakers, they have considerably less experience with language than adults, in the sense that they have had many fewer years of using language and have, for example, smaller vocabularies (see Goulden et al., 1990; Nagy et al., 1985; Zareva et al., 2005 for estimates of the adult active working lexicon; see Shipley & McAfee, 2015 for child vocabulary estimates). Informational cascades arise when there is rapid processing that results in quick information transfer across levels of representation. In the case of lexical processing, these interacting representations (word forms and meanings) must be acquired on the basis of experience. Thus, it seems plausible that considerable experience might be required before processing becomes efficient enough to support this level of incrementality. Indeed, it has been suggested that both the strength of the mappings between nodes and the interactivity of the lexical network increases with age (Bjorklund, 1995).

Domain-general limitations in children's cognition could also limit or prevent cascaded processing. For instance, five-year-olds have more limited working memory than adults (Chi, 1978; Cowan, 2017; Cowan et al., 2006; Dempster, 1981; Gathercole et al., 2004; Riggs et al., 2006; Schneider & Bjorklund, 1998; Simmering, 2012; inter alia). If working memory places limits on children's ability to simultaneously activate multiple lexical representations and multiple phonological forms, this could result in less cascading activation in younger children. Young children also have much weaker inhibitory abilities than adults: for example, they tend to be more susceptible to interference in Stroop tasks and other tasks requiring the suppression of a distractor (Bjorklund & Harnishfeger, 1990; Carter et al., 1995; Comalli et al., 1962; Dempster, 1992; Guttentag & Haith, 1978; Jerger et al., 2002; Jerger et al., 1988; Jerger et al., 1999; Ridderinkhof, 2002; Vurpillot & Ball, 1979). If multiple lexical and phonological representations are active at one time (as would occur within a cascading architecture), and children are unable to effectively inhibit candidates as new information comes in, then the costs of allowing activation to cascade across levels might be greater than the benefits. If this is the case, and if the processing system is adaptive (shaped by performance, over evolutionary or developmental time), then informational cascades might be suppressed until inhibitory abilities are more mature.

On the other hand, there are also reasons to think that one might find stronger or more widespread cascading activation in five-year-old children. First, if cascading activation is a basic property of cognitive architecture that cannot be suppressed (even when it is counterproductive) and if children have reduced inhibitory ability (see above), then the cognitive signatures of interactivity may be easier to detect in this population. A young child's inability to inhibit alternative lexical candidates might lead these candidates to be active for longer, allowing more time for information to cascade to the phonological level, providing a stronger signal of the cascade in measurements such as reaction times and speech errors. In contrast, older children and adults may more efficiently suppress the activation of non-target forms. This hypothesis could explain why Jescheniak and colleagues (2006) found evidence of semantically-mediated phonological activation in children seven to eight years old but not in older children or adults (Jescheniak et al., 2003; Jescheniak et al., 2006).

Second, young children's slower responses could result in more time for information to cascade prior to production. Five-year-old children typically take about 200 ms longer to name a picture than adults do (e.g. D'Amico et al., 2001 for five to six-year-olds; Jerger et al., 2002 for five to seven-year-olds). If this slowdown results from longer lexical selection times (as children consider competing lexical candidates), this additional time could allow more activation to cascade from the active lexical representations and their phonological forms, strengthening the activation of these forms beyond the levels they would typically reach in adults (assuming that the processes leading to the spread are not also slowed down). The effect of this slowdown thus might be similar to the effect of reduced inhibition described above. This increased activation of phonological forms could allow for more easy detection of cascaded activation, such as through semantically-mediated phonological effects.

By looking for signatures of cascading activation in five-year-old children, we can begin to target the questions of when and how this capacity emerges.

The present study

In the present study, we perform a side-by-side comparison of the mature and developing language production systems, comparing word planning by adults and five-year-old children in a picture naming task. We investigate the effects of image codability (name agreement) and name frequency on picture naming response time. These factors are commonly manipulated in language production experiments (e.g. Ferreira & Pashler, 2002; Griffin, 2001; Lee et al., 2013; Momma, 2021; inter alia), elicit robust effects across multiple languages and age groups (e.g. Bates et al., 2003; D'Amico et al., 2001; Johnson, 1992), and their effects are believed to index psychological processes within individuals during word planning. Crucially, they have been argued to influence the processes of lexical selection and phonological encoding, respectively, thereby allowing us to tap into the interplay between these processes in our two populations.

The codability of a referent serves as a measure of name agreement (e.g. Snodgrass & Vanderwart, 1980). Highly codable referents have a high level of name agreement (i.e. fewer alternative names applied to them), whereas referents with low codability have less name agreement (i.e. more possible names that can describe them). Speakers are faster to name pictures with higher codability than those with lower codability. Codability effects have been observed in both adults (e.g. Lachman, 1973; Lachman et al., 1974; Lachman & Lachman, 1980; Paivio et al., 1989) and in children, as young as four years of age (e.g. Butterfield & Butterfield, 1977; Johnson, 1992; Johnson & Clark, 1988). The codability effect is generally thought to influence the process of lexical selection, rather than visual processing or picture identification (Alario et al., 2004; Griffin, 2001; Johnson, 1992; inter alia). When there are multiple possible names for a speaker to choose from, these alternatives are simultaneously activated and compete with each other for selection. The greater the number of alternatives, the more competition, and thus the longer it takes for the speaker to resolve this competition and select a label. Recent research by Balatsou et al. (2022) supports this interpretation, finding evidence that name agreement is predictive of lexical co-activation within-speakers (though population-level measures may overestimate within-speaker variability).[2]

Word frequency has also been found to influence picture naming time in both adults (e.g. Bates et al., 2003; D'Amico et al., 2001; Jescheniak & Levelt, 1994; Lachman, 1973; Lachman et al., 1974; Oldfield & Wingfield, 1965; inter alia) and children (e.g. D'Amico et al., 2001 for five to six-year-olds). The frequency effect in word production is generally viewed as operating on the level of phonological encoding (e.g. Jescheniak & Levelt, 1994; see Griffin & Bock, 1998 for an overview of evidence in favour of this interpretation from RT and speech error studies; see Bates et al., 2003 for additional discussion of the loci of frequency RT effects), with more frequent word forms having higher resting activation, thereby allowing them to be accessed and selected more quickly than low frequency phonological forms during encoding. There is, however, some evidence that frequency may influence the lexical selection process in addition to affecting the selection of phonological form (e.g. Finocchiaro & Caramazza, 2006; Jescheniak & Levelt, 1994; Johnson et al., 1996; Kittredge et al., 2008; Strijkers et al., 2010), highlighting the complexity of attributing variables to specific processes within language production. Nevertheless, the effect of word form frequency is independent from that of conceptual frequency (Bates et al., 2003), and the effect of frequency on phonological encoding appears to be stronger and more reliable than its effect on lexical selection (Griffin & Bock, 1998).

Given that the codability effect has its locus in lexical selection and the frequency effect largely impacts phonological encoding, investigating how these two factors influence picture naming times allows us to investigate how these processes relate to one another during production. A statistical interaction between these effects could suggest that these processes interact, as is predicted by a cascading model of lexical planning. In contrast, factors that influence discrete, serial processing stages predict additive effects (Sternberg, 1969, 2001; but cf. Stafford & Gurney, 2011; Thomas, 2006). To our knowledge, such an interaction has not been previously reported for either child or adult picture naming. Critically, this is not because researchers have looked for interaction and failed to find it. The few picture naming studies we have found that directly explore both factors simply do not report on the presence or absence of an interaction in codability and frequency's influence on naming timing (e.g. Alario et al., 2004; Bates et al., 2003; Cycowicz et al., 1997; D'Amico et al., 2001; Lachman et al., 1974).[3]

Study 1 provides a side-by-side comparison of codability and frequency picture naming effects in five-year-olds and adults, assessing the similarity between adult and child lexical selection and phonological encoding processes. We replicate previously-observed RT effects of codability and frequency in our two populations. Critically, we also find an under-additive interaction between these effects. This pattern could suggest that phonological encoding begins before lexical selection ends, reducing the cost when both processes are slow – exactly what we would expect if there is an incremental cascade of information across these levels of representation. To further explore the effects of codability and frequency in our two populations, we fit ex-Gaussian distributions to our data to investigate how our manipulations influence the RT distributions (e.g. Staub, 2010), testing whether codability and frequency have qualitatively similar or different RT effects from each other and whether they influence adult and child participants similarly, suggesting comparable underlying processes. Study 2 tests the reliability of the interaction effect observed in Study 1, looking for comparable effects across languages in the adult naming data collected by Bates et al. (2003). Study 3 investigates how such an interaction might arise by simulating how the relationship between lexical selection and phonological encoding can influence RT in both serial and cascading activation architectures.

Data availability

The data, analysis code, and Supplementary Materials for all studies in this paper are available from: https://osf.io/myrtg/.

Study 1: a picture naming experiment with adults and five-year-old children

Study 1 investigated image codability and name frequency effects on picture naming RT in adults and five-year-old children. There were two sets of questions we sought to answer in this study:

- *Q1 (RT effects of codability and frequency)*: Are five-year-olds, like adults, slower to name pictures with low name agreement and low name frequency? Do the effects of codability and frequency interact in both populations?
- *Q2 (Influences of codability and frequency on the RT distribution)*: Do the codability and frequency manipulations influence RT distributions similarly? Are these RT distribution patterns qualitatively similar for adult and child responses (suggesting similar underlying processes) or different (suggesting changes in the lexicon over the course of development)?

Methods

The experiments in this study were approved by the Harvard University-Area Committee on the Use of Human Subjects (Protocol # 12718). The experiment methods were preregistered on OSF for both our adult (https://osf.io/hwtzs) and child (https://osf.io/3zcp8/) participants. The categorical analysis of codability and frequency in Q1 was preregistered. The remaining analyses in this paper were exploratory.

Participants

The participants were 48 adults (*M* age = 19.9 years, SD = 1.3; range = 18–23 years; 38 F, 10 M) and 25 children (*M* age = 5.47 years, SD = 0.29; range = 5;0–5;11; 7 F, 18 M). The number of child participants was determined via power analysis based on the effect sizes in the adult categorical analysis.[4] All participants were native, monolingual American English speakers. Adults were recruited from undergraduate classes at Harvard University and received partial course credit for their participation. Adult participants provided informed written consent to participate in the study. Children were recruited from the Harvard Laboratory for Developmental Studies database; they were given a small toy for participating, and their parents were given a $5.00 travel reimbursement. Informed written consent was received from the parent or guardian of the child participants for the child's participation. Six additional adults were tested but excluded from the analysis due to technical errors (1) or because they were early bilinguals (5). Three additional children were tested but excluded due to trial loss of over 50% (2) or because they were bilingual (1).

Materials

Participants viewed stimulus images from the BCBL MultiPic databank (Duñabeitia et al., 2018) and the Snodgrass and Vanderwart "Like" Objects (Rossion & Pourtois, 2004). The images were colourized digital images with black outlines. Adults saw and named 200 pictures. For the child experiment, we reduced the number of pictures to 120 so that the children would be more likely to complete the experiment. When selecting the images for the child experiment, we excluded items that received a large number of non-synonymous name responses in the adult data set, suggesting that the image was difficult to identify. For ease of comparison, the present analyses include only the responses to the 120 images named by both children and adults. We identify places where the result patterns differed in the full set of adult responses.

The experiment had a 2 × 2 within-subjects manipulation of image codability (high, low) and image name frequency (high, low), resulting in four conditions: *High Codability, High Frequency* (e.g. apple), *High Codability, Low Frequency* (e.g. cactus), *Low Codability, High Frequency* (e.g. sofa/couch), and *Low Codability, Low Frequency* (e.g. spaceship/UFO). To ensure that we could get responses that varied along the codability and frequency dimensions, for the adult stimulus set, we selected 50 images that we expected to fall into each of the four quadrants of our design (see Supplementary

Materials for details). None of the selected items were intended to elicit names that were conceptually or grammatically plural. After collecting the adult data (but prior to analysis), images were assigned to codability and frequency categories (see *Q1 data analysis procedure*). The 120 images named by both adult and child participants included 30 *High Codability, High Frequency* items, 29 *High Codability, Low Frequency* items, 29 *Low Codability, High Frequency* items, and 32 *Low Codability, Low Frequency* items.

Procedure

Each participant was tested individually in a single 20–30 min session. The stimulus images were shown on a Tobii T-60 remote eye-tracker. The session began with four practice trials (always in the same order) followed by the experimental trials, which were randomised.

At the beginning of each trial, a fixation cross appeared in the centre of the screen for 500 ms and was then replaced by a single stimulus image against a white background. For adult participants, the stimulus image remained on screen for four seconds before the trial ended. This procedure was slightly modified for the child participants. Given that similarly-aged children have been found to produce picture names more slowly than adults in comparable paradigms (e.g. D'Amico et al., 2001 for five- and six-year-olds), the maximum response time was increased to five seconds. As children are prone to loss of attention and/or fatigue in longer experiments, in order to minimise pauses and reduce the overall experiment duration, the experimenter advanced to the next trial after the participant named the object or indicated that they did not know the name.

Participants were instructed to name each image as quickly and as accurately as possible using a single word, to speak clearly, and to avoid producing any other words (e.g. articles) or sounds (e.g. clearing of the throat or filler sounds such as "umm") before giving a name. Audio responses were recorded through a microphone headset worn by the participant. The recording for each trial started when the stimulus image appeared on the screen and stopped when the image disappeared. Onset latencies were determined from these recordings.

Data exclusion & calculating onset

Responses were excluded from the analysis if the participant did not speak or did not name the object, if the audio recording did not contain the complete response, if the response contained more than a single name (renaming), if the response contained a false-start or repeated the start of the name, if the response contained a prenominal verbalisation (e.g. an article, "that's a … ", etc.),[5] if the response contained a prenominal sound (e.g. clearing of the throat, cough, "ummm", speech from the experimenter, participant, or parent) preventing the determination of the name onset time by the forced-aligner (see below), or if the onset time was otherwise incalculable from the recording (e.g. due to poor audio quality or background noise). We added an additional exclusion criterion beyond those specified in our preregistration to omit responses with post-nominal descriptions (e.g. "apple with a leaf", "baby touching its toes"), though we allowed responses in the form *N of N* ("ball of yarn", "jar of honey", "spool of thread", "loaf of bread"). Multi-word responses were excluded from the continuous analyses, as they lacked SUBTLEX-US frequency measures (Brysbaert & New, 2009) (see *Q1 data analysis procedure*).

For each response, the name was transcribed and speech onset time (measured from image onset) was determined using the Montreal Forced Aligner v1.0.0 (McAuliffe et al., 2017). Responses with prenominal sounds were flagged during the transcription process, and their alignments were checked; if onset time was identified by the forced aligner as the onset of the prenominal sound rather than the onset of the name, the response was omitted from analysis.

Q1 RT effects of codability and frequency

The first question we sought to answer in our experiment was whether codability and frequency influence naming RT in both the adult and child responses. We looked for previously-observed slowdowns in cases of low codability and low frequency, and we also looked for an interaction between the two effects.

Q1 data analysis procedure

All statistical analyses reported in the present paper were performed in R v 4.1.0 (R Core Team, 2021). Adult and child data were analysed separately.

Categorical analysis. Items were categorised into codability and frequency categories based on their name agreement and the frequency of the dominant names applied to them (their "target names") in the adult responses. When computing item name agreement (codability) and target names, we included all responses that gave a single complete label to the image, even if that response was not ultimately included in the analysis (e.g. responses including a determiner before the name; see *Data exclusion and calculating onset*).

An item's codability category (high, low) was determined based on its H score in the adult data. The H statistic is a measurement of name agreement calculated using the formula $H = \sum_{i=1}^{k} p_i \log_2 (1/p_i)$, where k is the number of different names given to the image, and p_i is the proportion of participants providing a specific name (e.g. Snodgrass & Vanderwart, 1980). The lowest possible value of H is 0, indicating perfect name agreement. The maximum H is achieved when each participant gives a different response and varies based on the number of participants in the experiment. Items were categorised based on the ratio between each image's H score and the maximum possible score in the experiment. The proportion of participants giving each name in the H score calculation was determined based on the number of responses included in the computation (rather than the total number of participants in the experiment). Items with H score ratios less than or equal to 0.06 were assigned to the high codability category, and items with H score ratios greater than 0.06 were assigned to the low codability category. The high codability items ($n = 59$) had an average H score of 0.07 (SD = 0.11) (M H score ratio = 0.01, SD = 0.02), and the low codability items ($n = 61$) had an average H score of 1.39 (SD = 0.57) (M H score ratio = 0.25, SD = 0.10).

An item's frequency category (high, low) was determined based on the frequency of its target name. Adult target names were used as the item target names for both the adult and child analyses. A frequency score was calculated for each item as a natural log transformation of its raw frequency in the SUBTLEX-US corpus (Brysbaert & New, 2009) [ln(1 + *raw frequency*), where *raw frequency* = words per million]. Items whose target names had frequency scores greater than 3.00 were assigned to the high frequency category, and items whose target names had frequency scores less than or equal to 3.00 were assigned to the low frequency category. The high frequency items ($n = 59$) had an average frequency score of 4.14 (SD = 0.90), and the low frequency items ($n = 61$) had an average frequency score of 1.74 (SD = 0.63).

We categorised the items based on adult measures because these measures are based on a larger number of responses (and are thus potentially more stable than corresponding child measures), and using the same categories for both populations allows for a direct comparison of their naming RT in response to the same items in the categorical analysis. Furthermore, the adult codability and frequency measures used for item categorisation are significantly correlated with the corresponding measures from the child data. Specifically, the H scores from the adult experiment ("adult H scores") and the H scores from the child experiment ("child H scores") had a Pearson's r of 0.67 ($t(118) = 9.79$, $p < 0.0001$). Similarly, there was a robust correlation between the items' frequency scores obtained from SUBTLEX-US ("frequency scores") and the frequency of their target names in the utterances of children aged 36 months and older the CHILDES corpus (MacWhinney, 2000) ("child frequency scores") (Pearson's $r = 0.81$, $t(118) = 14.96$, $p < 0.0001$).

The items in the high and low codability and frequency categories varied only along the dimensions we intended to manipulate. We confirmed this by conducting Type II Analyses of Variances (ANOVA) (performed using the R package {car} v3.0-11; Fox & Weisberg, 2019) and post-hoc pairwise-comparison tests (performed using {emmeans} v1.6.2-1; Lenth, 2019). The high and low adult codability items differed significantly based on adult H score ($t(118) = -17.42$, $p < 0.0001$) but did not differ significantly based on target name frequency score ($t(118) = 0.45$, $p = 0.66$). The high and low frequency items differed significantly based on target name frequency score ($t(118) = 17.06$, $p < 0.0001$) but not on adult H score ($t(118) = -1.21$, $p = 0.23$).

Table 1 shows the codability and frequency measures of our items broken down by the four experiment conditions (for additional properties of the items and their correlations with these measures, see Supplementary Materials). Critically, we succeeded in orthogonally manipulating codability and frequency across the four cells of our 2 × 2 design (confirmed via post-hoc pairwise-comparison tests with Tukey p-value adjustment; a full table of the pairwise comparison results is available in the Supplementary Materials). All of our codability and frequency measures reliably vary by condition ($F(3) \geq 35.80$, p's < 0.0001). We found significant differences in adult H score only between conditions with different codability categories ($|t(116)| \geq 10.95$, p's < 0.0001), and we found significant differences in adult frequency

Table 1. Properties of the codability × frequency conditions.

Condition	Adult H score	Child H score	Frequency score	Child frequency score
High Codability, High Frequency	0.02 (0.06)	0.20 (0.31)	4.34 (0.95)	6.17 (1.27)
High Codability, Low Frequency	0.11 (0.13)	1.07 (0.81)	1.58 (0.57)	2.39 (0.68)
Low Codability, High Frequency	1.30 (0.59)	1.51 (0.61)	3.94 (0.80)	5.40 (0.89)
Low Codability, Low Frequency	1.47 (0.56)	1.76 (0.69)	1.89 (0.64)	2.82 (1.28)

Mean adult H score, child H score, adult frequency score, and child frequency score for item target names in the four codability × frequency conditions. SD in parentheses.

score only between conditions with different frequency categories ($|t(116)| \geq 10.56$, p's < 0.0001).

The pattern of values for the child measures was more complicated. The child measurements differed significantly and substantially along the dimension we sought to manipulate: child H score differed between conditions with different codability categories ($|t(116)| \geq 2.66$, p's < 0.05), and child frequency scores differed between conditions with different frequency categories ($|t(116)| \geq 9.43$, p's < 0.0001). We did find, however, that the child H score also differed between the *High Codability, High Frequency* and *High Codability, Low Frequency* conditions, with higher child H scores (less name agreement) when frequency was low ($t(116) = -5.28$, $p < 0.0001$). In addition, the child frequency scores differed between the *High Codability, High Frequency* and *Low Codability, High Frequency* conditions, with higher child frequency scores in the high codability group ($t(116) = 2.75$, $p = 0.03$). These imbalances follow the direction of a previously-observed relationship between codability and frequency: name agreement tends to be higher for items with high frequency names (e.g. Bates et al., 2003). This pattern of differences could potentially result in an over-additive interaction in the child data analysis, as the *High Codability, High Frequency* is higher on the child measures of both of our manipulated dimensions than the other conditions. We address this concern in the continuous analysis.

We performed linear mixed effects analyses on log-transformed onset time (in log milliseconds) using the {lmerTest} package v3.1-3 (Kuznetsova et al., 2017). The linear mixed effects models used in the categorical analyses contained fixed effects for codability category (high, low), frequency category (high, low), trial number, and target name syllable count, with an interaction between frequency category and codability category and random slopes for codability category, frequency category, and their interaction by participant as well as a random intercept by item.[6] The fixed effects of trial number and target name syllable count were intended to control for effects of fatigue (e.g. D'Amico et al., 2001) and word length (e.g. D'Amico et al., 2001; Johnson et al., 1996; Székely et al., 2003; Székely et al., 2005; see Bates et al., 2003 for cross-linguistic differences in length effects) that increase picture naming RT. As the fixed effects of codability category and frequency category were entered into our analysis models with an interaction, to estimate the overall influence of these variables on RT, we used effects coding for these variables in the regression (e.g. Hardy, 1993), allowing us to compare the influence of the variables on the grand mean RT (analogous to main effects). To compute pairwise comparisons across levels of the categorical variables, we repeated the analyses with the same model structures using dummy-coding of the codability and frequency variables.

Continuous analysis. We addressed questions that arose from the results of the categorical analysis in an exploratory analysis (the "continuous analysis") that we had not pre-registered. To minimise the likelihood of false positives, wherever possible, we constrained the continuous analysis based on the decisions we made in the categorical analysis.

In this analysis, we investigated whether the data patterns observed in the categorical analysis persisted when we used continuous measures of codability and frequency. This analysis was intended to allay concerns that data patterns observed in the categorical analyses may be due to any departures from a perfectly orthogonal manipulation (particularly for the child data set). Paralleling the reported categorical results, we computed linear mixed effects models using the {lmerTest} package v3.1-3 (Kuznetsova et al., 2017) to analyse log-transformed onset time (in log milliseconds). When possible, the linear mixed effects models for the continuous analysis had the same effects structure as the categorical analysis, except that codability category and frequency category were replaced with H score and frequency score (any deviations from this structure are noted in the results). For maximum precision, we used the syllable counts and frequency scores of the individual responses rather than the items' target names. The adult analysis used adult H scores. We analysed the child data using both child and adult H scores; although the adult H scores reflect the name variation in the language input that children hear and were derived from a greater number of responses (as mentioned above), the child H scores may be a more direct reflection of child naming behaviour. Given the correlation between the frequency scores from SUBTLEX-US and from CHILDES (see above), we used the SUBTLEX-US frequency scores for both analyses, as they are based on a larger corpus of utterances.

Q1 results

Participant responses. For the 120 stimulus items that were viewed by both adult and child participants, we recorded 5760 adult responses and 3000 child responses. 74 adult responses and 593 child responses were excluded from the analyses based on the criteria above (see *Data exclusion & calculating onset*). The

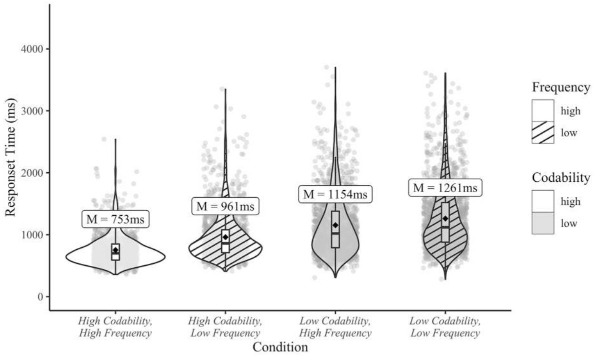

Figure 1. Adult RTs by condition. Each point represents a response. The black diamonds indicate mean RT.

distributions of false start, repeated start, renaming, and no response errors by codability and frequency category in the adult and child data are available in the Supplementary Materials. An additional 66 adult responses and 70 child responses were omitted from the continuous analyses because frequency counts were not available for them in SUBLTEX-US.

Replicating the results of previous studies (e.g. D'Amico et al., 2001), our child participants were in general slower to name the stimulus images than our adult participants: the average adult RT was 1034 ms (SD = 471 ms), and the average child RT was 1319 ms (SD = 632 ms). The distribution of onset times by condition are given in Figure 1 (adult data) and Figure 2 (child data).

Adult results. The categorical analysis of the adult data revealed a significant overall main effect of codability category ($\beta = -0.16$, $t(126) = -10.30$, $p < 0.0001$), with longer onset latencies for the low codability items. The estimated RT for low codability items (back-transformed from the log scale) was 1119ms (95% Confidence Interval [CI] [1050, 1192]), and the estimated RT for high codability items was 809ms (95% CI [759, 861]). We also observed a significant effect of frequency category ($\beta = -0.08$, $t(134) = -4.32$, $p < 0.0001$), with longer latencies for the low frequency items. The estimated RT for low frequency items was 1028ms (95% CI [965, 1095]), and the estimated RT for high frequency items was 882ms (95% CI [824, 944]). Effect plots showing the marginal effects of codability and frequency in the adult and child data are available in the Supplementary Materials. The main effect of target name syllable count was not significant in the present analysis ($\beta = 0.01$, $t(115) = 0.46$, $p = 0.65$), though the effect was significant in the full set of adult responses ($\beta = 0.04$, $t(5775) = 5.63$, $p < 0.0001$), with longer RTs for longer target names. The main effect of trial number was significant ($\beta = 2.4\text{e-}04$, $t(5491) = 3.86$, $p = 0.0001$), with longer RTs for later trials.

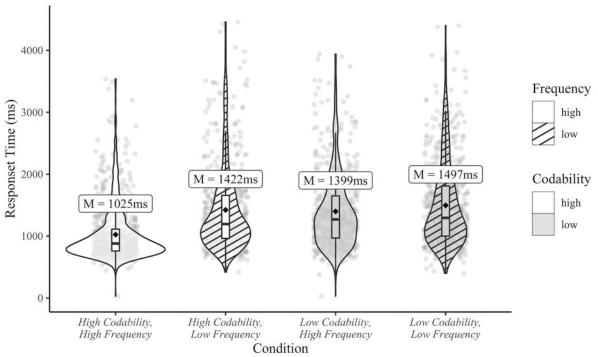

Figure 2. Child RTs by condition. Each point represents a response. The black diamonds indicate mean RT.

Crucially, there was a significant interaction between codability category and frequency category ($\beta = -0.03$, $t(119) = -2.17$, $p = 0.03$) such that the difference between high and low frequency was smaller for the low codability items than the high codability items (Figure 3). The effect of frequency category was significant in the high codability categories ($\beta = 0.22$, $t(134) = 4.47$, $p < 0.0001$) but not the low codability categories ($\beta = 0.09$, $t(120) = 1.92$, $p = 0.06$).

We observed the same pattern of results in the continuous analysis as in the categorical analysis (see Supplementary Materials for complete result summary). There were significant effects of item adult H score ($\beta = 0.20$, $t(193) = 8.51$, $p < 0.0001$) and response frequency score ($\beta = -0.05$, $t(269) = -4.63$, $p < 0.0001$), with slower RTs when H score was higher (i.e. indicating lower codability) and when frequency score was lower. Crucially, we also observed a significant interaction between adult H score and frequency score ($\beta = 0.02$, $t(205) = 2.59$, $p = 0.01$) that parallels the interaction between codability and frequency category observed in the categorical analysis: as H score increased (i.e. codability decreased), so did the effect of frequency score, meaning that the frequency effect was smaller for items with less name agreement (Figure 3).

Child results. The categorical analysis of the child data revealed significant overall main effects of codability category ($\beta = -0.09$, $t(86) = -5.24$, $p < 0.0001$) and frequency category ($\beta = -0.08$, $t(99) = -4.63$, $p < 0.0001$), with significantly longer RTs for the low codability and low frequency items. The estimated RT for low codability items was 1340 ms (95% CI [1258, 1427]), and the estimated RT for high codability items was 1113ms (95% CI [1046, 1184]). The estimated RT for low frequency items was 1330ms (95% CI [1247, 1418]), and the estimated RT for high frequency items was 1128ms (95%

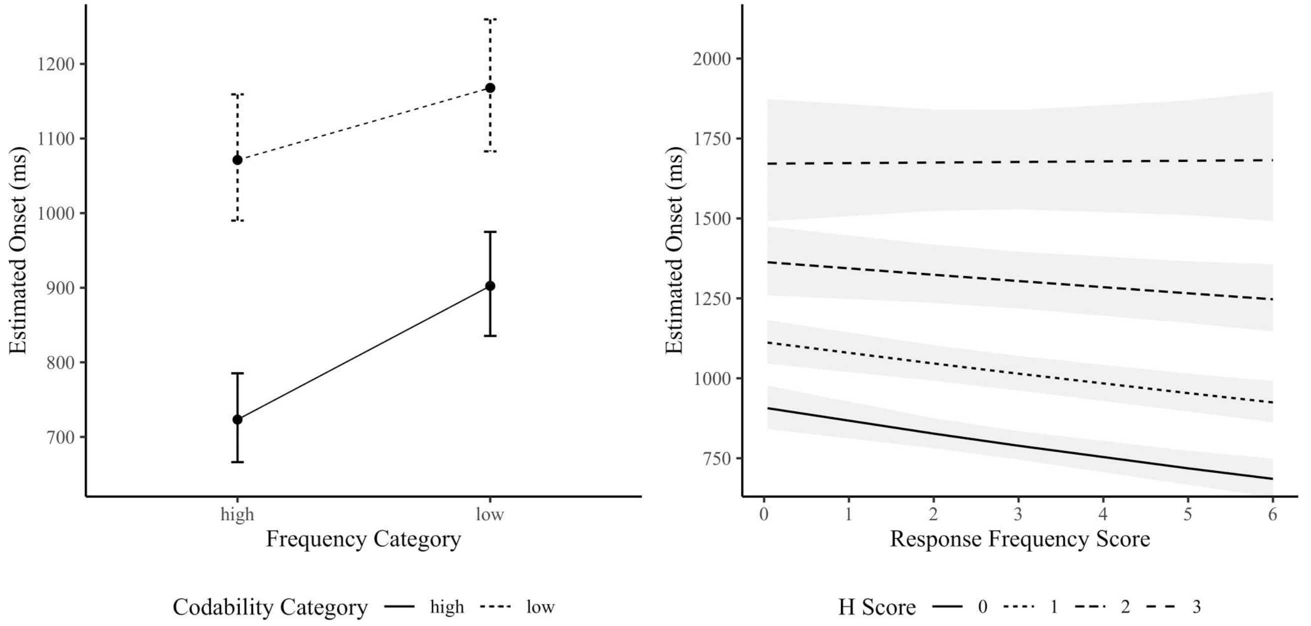

Figure 3. Interactions between codability and frequency measures in the adult data. RT estimates have been back-transformed from the analysis scale to the response scale (ms). Error bars and ribbons indicate 95% confidence intervals.

CI [1060, 1201]). The main effects of trial number ($\beta = 9.6\text{e-}04$, $t(2293) = 4.47$, $p < 0.0001$) and target name syllable count ($\beta = 0.06$, $t(113) = 2.18$, $p = 0.03$) were significant, with longer RTs for later trials and for items with longer target names.

Critically, as in the analysis of the adult data, there was a significant interaction between codability category and frequency category ($\beta = -0.06$, $t(105) = -3.83$, $p < 0.001$) such that the effect of frequency was larger when codability was high (Figure 4). The effect of frequency category was significant in the high codability categories ($\beta = 0.28$, $t(96) = 5.80$, $p < 0.0001$) but not in the low codability categories ($\beta = 0.04$, $t(109) = 0.90$, $p = 0.37$).

The same general pattern was present in the continuous analyses (complete results available in the Supplementary Materials). For the analysis using child H score as our continuous measure of codability, we had to drop the interaction term from the participant random effect for model convergence. In the resulting model, we observed significant effects of child H score ($\beta = 0.13$, $t(182) = 4.58$, $p < 0.0001$) and response frequency score ($\beta = -0.05$, $t(128) = -4.08$, $p < 0.0001$), and a significant interaction between the two variables ($\beta = 0.02$, $t(281) = 2.68$, $p < 0.01$) (Figure 4). In the analysis using adult H score rather than the child H score, we were able to use the full random effects structure. In this model, we observed a significant interaction between adult H score and response frequency score ($\beta = 0.02$, $t(56) = 2.04$, $p < 0.05$) and a significant effect of response frequency score ($\beta = -0.06$, $t(160) = -4.59$, $p < 0.0001$). The marginal effect of adult H score did not reach conventional levels of significance with the full random effects structure ($\beta = 0.08$, $t(65) = 1.97$, $p = 0.05$) but did in a model with the same random effects structure as the child H score analysis ($\beta = 0.08$, $t(262) = 2.09$, $p = 0.04$).

Q1 summary

In the Q1 analyses, we replicated previously-observed codability and frequency effects in both adults and five-year-old children: naming RT was faster when images had high codability (more name agreement) and when their names were more frequent. In both populations, we additionally observed under-additive interactions between the effects of codability and frequency such that the frequency effect was attenuated when codability was low. These effects were observed using both categorical and continuous measures of codability and frequency. We address a possible explanation for these under-additive interactions in Study 3, suggesting that the interaction stems from the dynamics of the language production system.

One concern that is always present in reaction time studies is that timing differences across conditions could be a side effect of differences in the proportion of responses that were discarded. For example, we omitted responses that were not completed within the response window; if these omitted responses were particularly common for items with low codability and low

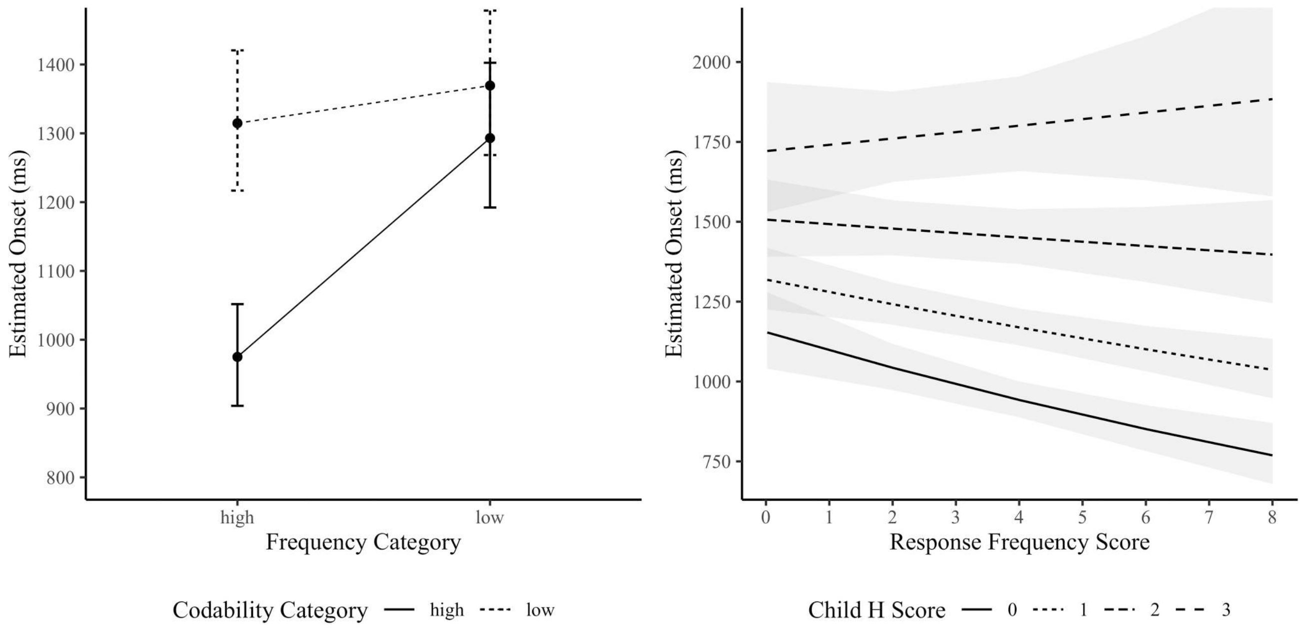

Figure 4. Interactions between codability and frequency measures in the child data. RT estimates have been back-transformed from the analysis scale to the response scale (ms). Error bars and ribbons indicate 95% confidence intervals.

frequency (especially compared to high codability, low frequency items), then the absence of these longer response times could result in an artificial under-additive interaction. We saw no evidence for this pattern. Less than 0.5% of all responses were omitted for this reason in both the adult and child data sets, and these omissions were scattered across the four cells of the design (see Supplementary Materials). Similarly, speech errors resulting in response omission do not appear to be more common for responses to *Low Codability, Low Frequency* items compared to those to *High Codability, Low Frequency* items in either population (see Supplementary Materials). We consequently do not believe that the observed under-additive interactions are a by-product of response omissions.

In addition to observing effects of codability and frequency, our analyses also reproduced other established naming RT patterns in adults and children. Both populations displayed fatigue effects (e.g. D'Amico et al., 2001), with slower response times to later trials, though the fatigue effect was more pronounced in the child data. We also replicated effects of word length (e.g. Bates et al., 2003; D'Amico et al., 2001; Johnson et al., 1996; Székely et al., 2003; Székely et al., 2005), with faster RTs for shorter names. In a post-hoc exploratory analysis, we additionally assessed the influence of age of acquisition (AoA) on naming response time (see Supplementary Materials). Words rated as having been acquired earlier in life tend to be produced faster and with fewer errors (e.g. Carroll & White, 1973; see Brysbaert & Ghyselinck, 2007 for review; see Anderson, 2008; D'Amico et al., 2001; Johnson & Clark, 1988 for evidence of AoA effects in children). AoA was omitted from our preregistered analysis for two reasons. First, there is a strong correlation between frequency and AoA since more frequent words tend to be learned at earlier ages (see, e.g. Goodman et al., 2008), raising statistical concerns. Second, there is debate over whether AoA measures capture the same cognitive construct as word frequency measures (e.g. Zevin & Seidenberg, 2002). Our post-hoc analyses replicate previously-observed AoA effects in adults and children and provide evidence that AoA and frequency effects pattern differently in our data set, suggesting that frequency and AoA have independent effects (Brysbaert & Ghyselinck, 2007; Juhasz, 2005) and that the observed interaction between codability and frequency is not driven by AoA. In an additional exploratory analyses, we also confirm that the observed interaction between codability and frequency is independent from effects of phonological neighbourhood density (see Supplementary Materials).

Overall, we observed very similar results in both the adult and child analyses, suggesting similar underlying processes in both populations. We also found broadly similar results regardless of whether we based our predictors on the child data or the adult data. Nevertheless, in the continuous analysis, we observed a greater estimated effect on child naming RT for child H score (Cohen's $d = 0.68$) than adult H score (Cohen's $d = 0.49$ with the full effects structure, Cohen's $d = 0.26$ with the same effects structure as the child H model) (standardised effects sizes computed using {EMAtools} v0.1.4; Kleiman, 2021). These differences suggest that child H scores may serve as a more accurate reflection of child naming behaviour. Differences between adult and child H scores could reflect differences in the words that children know or in the contexts in which they use them, resulting in differences in adult and child name agreement.

The results of the Q1 analyses demonstrate that by five years of age, lexical production in children depends upon both the codability of the referent and the frequency of the word. These RT effects suggest that the process of lexical production is similar in both adults and five-year-olds, as both populations display codability and frequency effects on picture naming RT. However, mean differences in response time can arise through a variety of different changes in the response time distribution. These changes are thought to reflect different types of processing costs (some of which affect all trials, some of which affect only a subset). In the next section, we explore how our two factors (codability and frequency) affect the response time distribution in children and adults to assess whether the underlying processes affected by these manipulations are qualitatively similar in the two populations.

Q2: influences of codability and frequency on the response time distribution

In this section, we explore how the codability and frequency manipulations affected the RT distributions in the adult and child data. A variable can affect RT by shifting the mean (increasing the RT for all trials), changing the standard deviation (increasing the variability of RT), or skewing the distribution (increasing the RT for a subset of trials). These different changes are thought to reflect different underlying processes (Balota & Spieler, 1999). By analysing the RT distribution and how it changed for each manipulation in each population, we can gain insight into the processes involved in lexical production and how they change (or stay the same) between five years of age and adulthood.

Q2 data analysis procedure

To investigate how the frequency and codability manipulations impacted the RT distributions in our data, we fit ex-Gaussian distributions (Ratcliff, 1979) to the RT data from each participant in each codability and frequency category (high, low). Ex-Gaussian distributions are often used to describe RT distributions (Balota & Spieler, 1999; Dawson, 1988; Luce, 1986; Ratcliff, 1993). Ex-Gaussian distributions are convolutions of a normal distribution (described by parameters μ and σ) with an exponential distribution (described by the parameter τ). The way that a manipulation influences these parameters can serve as an indication of how it influences response time. If a manipulation shifts the mean of the RT distribution, it will primarily influence μ. Changes in the standard deviation of the distribution will influence σ. If a manipulation increases the skew of the RT distribution, it will primarily influence τ.

Parameter estimations in our analysis were computed using the maximum likelihood method. We fit the ex-Gaussian distributions in R using functions from the {retimes} v0.1-2 package (Massidda, 2013). We constructed linear mixed effects models (using {lmerTest} v3.1-3; Kuznetsova et al., 2017) to analyse the estimates obtained for each parameter in each category. The models had fixed effects of category level (high vs. low) and manipulation (codability, frequency) with an interaction, as well as a random intercept for participant. We used dummy-coding to establish the contrasts of interest.

We supported this analysis with vincentile plots (Vincent, 1912), which serve as a non-parametric confirmation of ex-Gaussian analyses (e.g. Balota et al., 2008; Staub, 2010). We constructed these plots by dividing the data for each participant in each condition into ten vincentiles (the fastest 10% of responses, the next fastest 10%, etc.). We then calculated the mean RT for each participant in each condition (high, low) in each vincentile as well as the difference between these means. The vincentile plots show the mean RT difference between conditions across all participants at each vincentile; a vincentile plot thus demonstrates how the size of an effect changes across the RT distribution. The shapes of vincentile plots systematically reflect effects on ex-Gaussian parameters (Balota et al., 2008). A manipulation that shifts a RT distribution, resulting in a change in μ, will have a relatively flat vincentile plot, with similar mean RT differences for all ten vincentiles. Changing the size of σ will leverage the plot around a midpoint, increasing slope as σ increases. A manipulation that skews the RT distribution, manifesting in a change in τ, will have greater mean RT differences in the tail of the distribution (i.e. at larger vincentiles), resulting in a vincentile plot that curves upwards as vincentile number increases.

We performed separate analyses for the adult and child data. For the adult data, we used the same codability and frequency categories as in the Q1 categorical analysis. For the child data, we reclassified the items' codability categories based on their child H scores, because, as we noted in the *Q1 Summary*, the child H scores appeared to better capture the codability effect in our child data than adult H scores did. The items were categorised using the same method described in the *Q1 Analysis Procedure* but with child H score ratios.

We refer to these new categories as the "child codability" categories.

Table 2 illustrates the adult and child codability and frequency measures for the categories in the analysis. As mentioned in the *Q1 analysis procedure*, the high and low frequency categories and the codability categories based on adult H score only vary along the dimension we intended to manipulate. The child codability categories, on the other hand, were less well-balanced. The high and low child codability items differed significantly based on both adult H score ($t(118) = -7.20$, $p < 0.0001$) and child H score ($t(118) = -12.50$, $p < 0.0001$), though they also differed significantly based on frequency score ($t(118) = 3.90$, $p < 0.001$). The high and low frequency items additionally differed based on child H score ($t(118) = -3.93$, $p = 0.0001$). These imbalances reflect the trend that name agreement tends to be higher for items with high frequency names (e.g. Bates et al., 2003). Given the imbalances in the child codability categories, we also analysed the child data using the adult codability categories (relevant differences will be discussed in the text; see Supplementary Materials for full results). We address potential influences of these imbalances in the *Q2 summary*.

Q2 results

The mean estimated values of the three ex-Gaussian parameters for adult and child data in each level of the codability and frequency manipulations are given in Table 3 along with the estimated RT differences between the levels. Reliability of the low – high difference for the ex-Gaussian parameters is indicated in the table.

Adult results. The μ estimates were larger in the low codability condition than the high codability condition ($\beta = 125.82$, $t(141) = 12.34$, $p < 0.0001$) and in the low frequency condition than the high frequency condition ($\beta = 86.96$, $t(141) = 8.53$, $p < 0.0001$). There was a significant interaction between level and manipulation such that the difference between the codability conditions was greater than the difference between the frequency conditions ($\beta = -38.86$, $t(141) = -2.69$, $p < 0.01$).

Table 2. Properties of the codability and frequency categories.

Manipulation	Category	Adult H score	Child H score	Frequency score
Codability (Adult)	High ($n = 59$)	0.07 (0.11)	0.63 (0.75)	2.98 (1.60)
	Low ($n = 61$)	1.39 (0.57)	1.64 (0.66)	2.87 (1.25)
Codability (Child)	High ($n = 33$)	0.04 (0.09)	0.08 (0.12)	3.70 (1.57)
	Low ($n = 87$)	1.00 (0.77)	1.54 (0.66)	2.63 (1.26)
Frequency	High ($n = 59$)	0.65 (0.76)	0.84 (0.82)	4.14 (0.90)
	Low ($n = 61$)	0.82 (0.80)	1.43 (0.82)	1.74 (0.63)

Mean adult H score, child H score, and frequency score for the high and low codability and frequency categories used in the Q2 analyses. SD in parentheses.

The σ estimates were larger in the low codability condition than the high codability condition ($\beta = 43.37$, $t(141) = 5.15$, $p < 0.0001$). There was no reliable difference between the low and high frequency groups in the present analysis ($\beta = 12.07$, $t(141) = 1.43$, $p = 0.15$), though estimates were reliably larger in the low frequency condition in the full adult data set ($\beta = 16.26$, $t(141) = 2.12$, $p = 0.04$). There was a significant interaction between level and manipulation such that the codability manipulation had a greater effect on σ than the frequency manipulation ($\beta = -31.30$, $t(141) = -2.63$, $p < 0.01$).

The τ estimates were larger in the low codability condition than the high codability condition ($\beta = 231.26$, $t(141) = 18.56$, $p < 0.0001$) and in the low frequency condition than the high frequency condition ($\beta = 83.04$, $t(141) = 6.67$, $p < 0.0001$). There was a significant interaction between level and manipulation such that the effect on τ was larger for the codability manipulation than the frequency manipulation ($\beta = -148.23$, $t(141) = -8.41$, $p < 0.0001$).

Figure 5 provides the vincentile plots of the codability and frequency manipulations in the adult data. The RT difference between high and low codability conditions increased from the faster to the slower RT vincentiles. The upward curve of the codability plot resembles the idealised plot for a RT effect due to a change in τ (exponential contribution) (Balota et al., 2008). The plotted line for the frequency manipulation, on the other hand, has a more linear shape, with the effect size increasing linearly across the RT distribution. This vincentile plot shape is consistent with those for RT effects that are divided between μ (mean) and τ (exponential contribution) (Balota et al., 2008).

Child results. In the child analysis, the μ estimates were larger in the low codability condition than the high codability condition ($\beta = 111.24$, $t(72) = 7.71$, $p < 0.0001$) and in the low frequency condition than the high frequency

Table 3. Mean RT and ex-Gaussian parameter estimates (in ms) for the high and low codability and frequency conditions.

	Adult data				Child data			
Condition	RT	μ	σ	τ	RT	μ	σ	τ
High codability	854	603	65	252	1035	688	53	353
Low codability	1210	729	108	483	1450	800	121	641
Low – High	356	126	43	231	415	112	63	288
Reliability		***	***	***		***	***	***
High frequency	948	585	56	364	1194	693	79	500
Low frequency	1117	672	68	447	1462	792	94	661
Low – High	169	87	12	83	268	99	15	161
Reliability		***	n.s.	***		***	n.s.	***

For the low – high difference reliability, "***" indicates p-values ≤ 0.001, and "n.s." indicates p-values > 0.1.

condition ($\beta = 99.12$, $t(72) = 6.87$, $p < 0.0001$). There was no significant interaction between level and manipulation for μ ($\beta = -12.13$, $t(72) = -0.59$, $p = 0.55$).

The σ estimates were larger in the low codability condition than the high codability condition ($\beta = 68.15$, $t(72) = 3.80$, $p < 0.001$). There was no reliable effect of frequency condition on the σ estimates ($\beta = 15.54$, $t(72) = 0.87$, $p = 0.39$). The interaction between level and manipulation was significant such that the codability manipulation had a greater effect on σ than the frequency manipulation ($\beta = -52.61$, $t(72) = -2.08$, $p = 0.04$), though this interaction did not reach conventional levels of significance in the analysis with adult codability categories ($\beta = -51.69$, $t(72) = -1.97$, $p = 0.05$).

The τ estimates were larger in the low codability condition than the high codability condition ($\beta = 288.02$, $t(72) = 10.81$, $p < 0.0001$) and in the low frequency condition than the high frequency condition ($\beta = 161.08$, $t(72) = 6.05$, $p < 0.0001$). The interaction between level and manipulation was significant such that the effect on τ was larger between the codability conditions than the frequency conditions ($\beta = -126.94$, $t(72) = -3.37$, $p = 0.001$). This interaction was not significant in the analysis using adult codability categories.

Figure 5 shows the codability and frequency vincentile plots for the child data. The codability plot shows a similar upward curve to that observed in the adult analysis, consistent with a large τ effect (skewing the RT distribution). The plotted line for the frequency manipulation is primarily linear, though it has a slight upward curve; this is consistent with the finding that the τ effect (exponential contribution) for frequency was larger than the μ effect (mean) though still smaller than the τ effect for child codability. The greater slope of the frequency plot compared to the adult frequency vincentile plot is consistent with the fact that children had a larger frequency RT effect (268 ms) than adults (169 ms).

Q2 summary

In the Q2 analyses, we assessed whether the Study 1 codability and frequency manipulations had qualitatively similar influences on adult and child naming RT, suggesting similar underlying processes. We used ex-Gaussian analyses to estimate how codability and frequency influenced the RT distributions for adults and children, investigating which parameters of the distribution were affected by the manipulations.

The ex-Gaussian analyses suggest that codability and frequency effects are qualitatively different from each other but similar across the two age groups. For both children and adults, the frequency manipulation had effects on the μ (mean) and τ (exponential contribution) parameters, suggesting that decreasing frequency both shifts and skews the RT distribution. The codability manipulation, on the other hand, influenced all three parameters in both populations, with a particularly prominent τ (skew) effect.[7] This skew effect was not only greater than codability's influence on the other two parameters, but it was also considerably larger than the τ effect of the frequency manipulation.

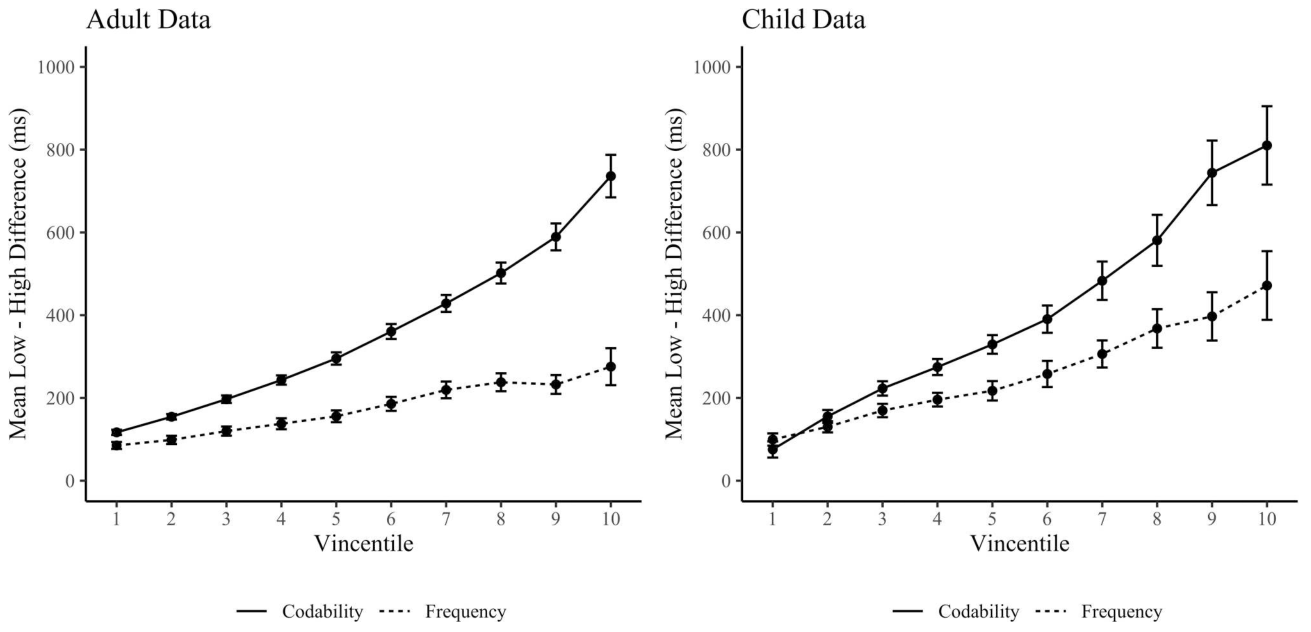

Figure 5. Vincentile plots. These plots illustrate the difference between the high and low conditions of each manipulation across the RT distribution. Error bars indicate standard errors.

The fact that the codability and frequency manipulations produced different effects on RT distribution suggests that the two variables play different roles in lexical processing. These findings are therefore consistent with hypotheses in which each factor affects a different part of the lexical access process. The ex-Gaussian parameters influenced by codability and frequency manipulations provide clues as to what these processes may be. Differences in the Gaussian parameters (μ and σ) are often interpreted as reflecting changes in automatic processes (Balota & Spieler, 1999), such as the initial perceptual processing or activation of candidates. In contrast, a large shift in the exponential parameter (τ) is the hallmark of decision-making processes (Hohle, 1965).

The large τ effect for the codability manipulation thus supports the interpretation of the codability effect as reflecting increased time to resolve competition and select between name candidates during lexical selection. This hypothesis can explain codability's influence on all three parameters. We should expect variation in RT slowdowns for low codability items, as there is variation in the number of candidate names available for each picture (including due to individual differences; e.g. one individual may always refer to a couch using the word *couch*, whereas another may use the words *sofa* and *couch* interchangeably) as well as in the relative frequencies of these names. These variations should result in larger penalties for some trials (e.g. when there are a greater number of names under consideration) and smaller or no influence for others (e.g. when there are fewer name candidates under consideration or one candidate with higher frequency than the others). By contrast, we expect more similar (and faster) RTs in the high codability condition when there are consistently few or no name alternatives to decide between and participants largely produce the same responses. This difference would lead to increased skewing in the low codability condition (a τ effect), and the σ and μ effects would arise as a consequence of these same forces: the variability in the low codability RTs in the low codability condition would increase the standard deviation around the mean relative to the high codability condition (leading to a σ effect), and the larger proportion of slow RTs in the low codability condition would result in a larger mean RT, shifting the distribution (resulting in a μ effect). The fact that codability manipulation influenced all three parameters could additionally reflect influences of the manipulation at multiple levels of processing (e.g. conceptual processing/image identification in addition to lexical selection).

On the other hand, a phonological frequency effect neatly accounts for the RT distribution changes we observed for the frequency manipulation. If the frequency effect reflects differences in the time to activate phonological forms, we should expect similar RT penalties for names with similar frequencies. Thus, we would expect a rightward shift of the RT distribution in the low frequency condition (a μ effect), as speakers are consistently slower to produce these names. Consistent with this account, the frequency manipulation produced a smaller skewing effect (τ effect) than the codability manipulation and no σ effect, demonstrating that slowing in the low frequency condition was more homogenous than in the low codability condition.

While frequency's τ effect was smaller than that for codability, it was still reliable. Skewing of the RT distribution in the low frequency condition could arise for several reasons, all of which are compatible with a phonological frequency effect: (i) Responses were sorted into high and low frequency categories based on the frequencies of the item dominant names, rather than the individual responses; this leaves open the possibility that these categories do not accurately reflect the frequencies of all the responses they contain. (ii) There is variation in the dominant name frequencies of the items categorised as low codability, which will result in greater RT penalties for some responses compared to others in the category. (iii) There is individual variation in the frequencies with which names are produced and/or encountered, and this variation is likely to be greater for low frequency words than for high frequency words. (iv) Increased skewing for low frequency items may arise due to lateral inhibition at the phonological level or due to the shape of the activation function and different activation thresholds required for selection (Andrews & Heathcote, 2001; Balota & Spieler, 1999). (v) Since item name agreement tends to pattern with name frequency (e.g. naming disparity is smaller for items eliciting high frequency names; Bates et al., 2003), the frequency manipulation may also be a weak manipulation of codability, which could contribute to the skewing effect.

In sum, although the ex-Gaussian analyses do not uniquely support a model of language production in which codability affects lexical selection and frequency affects phonological encoding, they are consistent with it. Minimally, these analyses show that codability and frequency manipulations influence RT distributions in ways that are different from each other.

Critically, the pattern of effects observed for the two manipulations was qualitatively similar in the two age groups investigated. This close parallelism between the children and adults suggests that the mature and developing lexical access systems involve similar underlying mechanisms that exhibit comparable responses to these factors. Nevertheless, there were some more subtle differences between the adult and child groups. The child population demonstrated a larger RT effect of the frequency manipulation than the adult population (268 ms vs. 169 ms, respectively). In adults, the frequency had similar effects on the mean shift (87 ms) and skew (83 ms), while in children the effect of frequency on skew was more pronounced (161 ms compared to a 99 ms shift effect). This difference may be attributable, in whole or part, to a weak correlation between codability and frequency measures that was present in the child data set but not the adult data set. Specifically, the child H scores for items in the low frequency condition were slightly lower than those for the high frequency items (Table 2). For this reason, the frequency manipulation in children is likely also a weak manipulation of codability, resulting in a pattern of effects that is intermediate between the adult frequency and codability effects. This manipulation of codability in the child data would result in an increase in the skew effect for frequency in children, relative to adults, which in turn could contribute to the larger average RT penalty for low frequency trials.

Study 1 discussion

In Study 1, we explored the effects of frequency and codability on lexical production in five-year-old children and adults. Our findings confirmed two previously-reported patterns: participants in both populations were faster to name images with higher codability (higher name agreement) and with higher frequency. In addition, we went beyond the prior work and fit ex-Gaussian distributions to the reaction time data. We found that codability and frequency manipulations have distinct influences on RT distributions, and these signature patterns are present in both the adult and child data. Taken together, these findings suggest that codability and frequency influence different underlying processes and that their effects are similar in the mature and developing lexical production systems, potentially indicative of similar underlying processes.

Critically, Study 1 found that codability and frequency effects interact in both adult and five-year-old naming behaviour: in both the adult and child data, the frequency effect was diminished when codability was low. To our knowledge, such an interaction has not been previously reported for either population. Indeed, early studies suggested that these effects are independent (Lachman, 1973; Lachman & Lachman, 1980), making the interactions observed in our data particularly unexpected. The observed interactions are noteworthy, because in standard models of reaction times, an interaction between two factors implies that these two variables influence at least one common process, either because both are inputs into that process or because they are inputs into separate processes that then interact (Sternberg, 1984, 1998). This is potentially surprising because codability and frequency are thought to influence distinct processes within word planning: the processes of lexical selection and phonological encoding, respectively. Thus, an interaction might suggest that these two processes do not occur in a strict sequence but instead influence one another, as is predicted by a cascading activation architecture.

In Study 2, we investigate the reliability of the observed interaction between codability and frequency in a secondary analysis of data from a previous naming study conducted with adults in several languages (Bates et al., 2003). In Study 3, we explore how such an interaction may arise based on the relationship between lexical selection and phonological encoding in the production planning architecture.

Study 2: a secondary analysis of prior adult picture naming data

The goal of Study 2 was to determine whether the interactions that we observed in Study 1 between codability and frequency could be observed in other data sets with different properties. Study 1 had two clear limitations. First, we explored naming in just one language (English). If our hypothesis is correct and the informational cascade is a foundational property of the linguistic architecture, then we should see this same under-additive interaction in other languages as well. Second, our experimental stimuli were initially selected to create four conditions that orthogonally manipulated our two variables, supporting a categorical analysis. As a result, the items did not represent a full spectrum of codability or frequency measures. Thus, it is possible that the interaction effect we observed is a side effect of stimulus selection under these constraints or is limited to extreme values of frequency and codability.

To investigate the generality of the interaction, we looked for parallel effects in the data from Bates et al.'s (2003) multi-language timed picture naming study. In this paper, the authors reported influences of name agreement and word frequency on mean RT in the expected directions (higher H scores and lower word frequencies predicted slower mean RTs), but they did not test for an interaction between codability and frequency RT effects. This data set complements ours because the stimuli were not selected specifically with a manipulation of codability and frequency in mind. Furthermore, it allows us to explore whether the interaction we observed is present across a range of languages.

Methods

We re-analysed the picture naming data from Bates et al.'s (2003) multi-language naming study, using the data set available from https://crl.ucsd.edu/experiments/ipnp/7lgpno.html. Our analysis focused on the data for adult native speakers of: English ($N = 50$), German ($N = 30$), Hungarian ($N = 50$), Italian ($N = 50$), Mandarin Chinese ($N = 50$), and Spanish ($N = 50$).[8] Participants named a set of 520 black-and-white images. For details about the stimulus materials and experiment methods used, please see Bates et al. (2003).

Analysis

The data set available from Bates et al. (2003) provides mean reaction times for each item in each language (rather than individual trial response data). To parallel our RT analyses in Study 1, we conducted linear regressions on log-transformed mean RT (in log milliseconds). We analysed the data from each language separately using regression models with fixed effects of item H score, dominant name frequency score, and dominant name syllable count, with an interaction between H score and frequency score. We used the dominant name frequency scores and H scores provided in the Bates et al. (2003) data set. The frequency scores in the data set were derived as natural log transformations from words per million corpus counts (as in Study 1). The properties of the codability and frequency measurements for each language are summarised in Table 4.

We also analysed the data from all six languages in a single model with fixed effects of item H score, dominant name frequency score, language, and dominant name syllable count, with a three-way interaction between H score, frequency score, and language.

Results

The results of our analyses are summarised in Table 5. All six languages analysed had significant fixed effects of H score and frequency score. Mean RTs increased as H scores increased (i.e. codability decreased) and as frequency score decreased. The interaction between H score and frequency score was significant in the English, German, Mandarin, and Spanish data: as H score increased, the effect of frequency score decreased. There was no statistically significant interaction in Hungarian or Italian. In the combined model looking at the data from all languages, we observed significant fixed effects of H score, frequency score, and their interaction (see Supplementary Materials for the complete model summary).

Plots showing the interaction between H score and frequency score in each analysis are presented in Figure 6. The plots for the languages with significant interactions between H score and frequency score (English, German, Mandarin, and Spanish) resemble those for the adult and child data in Study 1 (Figures 3 and 4): the slope of the estimated frequency effect becomes less negative at higher H score values (i.e. at lower levels of name agreement). We observe a similar pattern in the interaction plot for the combined language analysis. For the two languages without significant interactions (Hungarian and Italian), the slope of the estimated frequency effect is similar at each H score level.

Table 4. H score and frequency score measures for each language in Bates et al. (2003).

	H score		Frequency score	
Language	Mean	Range	Mean	Range
English	0.67 (SD = 0.61)	0.00–2.90	2.50 (SD = 1.57)	0.00–7.40
German	0.76 (SD = 0.68)	0.00–3.28	2.01 (SD = 1.50)	0.00–6.62
Hungarian	0.91 (SD = 0.73)	0.00–3.52	1.38 (SD = 1.93)	0.00–6.84
Italian	0.95 (SD = 0.73)	0.00–3.47	1.17 (SD = 1.43)	0.00–6.20
Mandarin	1.16 (SD = 0.79)	0.00–3.56	3.05 (SD = 1.65)	0.00–7.60
Spanish	0.86 (SD = 0.72)	0.00–2.90	2.77 (SD = 1.78)	0.00–8.32

Study 2 discussion

Study 2 demonstrates that the under-additive interaction between codability and frequency observed in Study 1 generalises across several languages and to a

Table 5. Model results for the re-analysis of Bates et al.'s (2003) multi-language naming data.

Language	H score	Frequency score	Syllable count	Interaction
English	$\beta = 0.18$ $t(515) = 9.11$ $p < 0.0001$	$\beta = -0.04$ $t(515) = -5.97$ $p < 0.0001$	$\beta = 2.7\text{e-}03$ $t(515) = 0.30$ $p = 0.76$	$\beta = 0.02$ $t(515) = 3.19$ $p < 0.01$
German	$\beta = 0.18$ $t(515) = 10.61$ $p < 0.0001$	$\beta = -0.05$ $t(515) = -6.25$ $p < 0.0001$	$\beta = 0.01$ $t(515) = 0.85$ $p = 0.40$	$\beta = 0.02$ $t(515) = 3.22$ $p = 0.001$
Hungarian	$\beta = 0.24$ $t(515) = 20.57$ $p < 0.0001$	$\beta = -0.02$ $t(515) = -3.46$ $p < 0.001$	$\beta = 2.2\text{e-}03$ $t(515) = 0.29$ $p = 0.77$	$\beta = 1.1\text{e-}04$ $t(515) = 0.02$ $p = 0.98$
Italian	$\beta = 0.21$ $t(515) = 18.20$ $p < 0.0001$	$\beta = -0.03$ $t(515) = -4.75$ $p < 0.0001$	$\beta = -0.01$ $t(515) = -1.04$ $p = 0.30$	$\beta = 1.3\text{e-}03$ $t(515) = 0.19$ $p = 0.85$
Mandarin	$\beta = 0.17$ $t(515) = 9.09$ $p < 0.0001$	$\beta = -0.05$ $t(515) = -5.61$ $p < 0.0001$	$\beta = 0.02$ $t(515) = 1.47$ $p = 0.14$	$\beta = 0.01$ $t(515) = 2.62$ $p < 0.01$
Spanish	$\beta = 0.20$ $t(515) = 12.58$ $p < 0.0001$	$\beta = -0.03$ $t(515) = -5.28$ $p < 0.0001$	$\beta = -8.1\text{-}04$ $t(515) = -0.11$ $p = 0.91$	$\beta = 0.01$ $t(515) = 2.69$ $p < 0.01$
Combined Language Model	$\beta = 0.18$ $t(3095) = 8.65$ $p < 0.0001$	$\beta = -0.04$ $t(3095) = -5.91$ $p < 0.0001$	$\beta = 1.7\text{e-}03$ $t(3095) = 0.48$ $p = 0.63$	$\beta = 0.02$ $t(3095) = 3.03$ $p < 0.01$

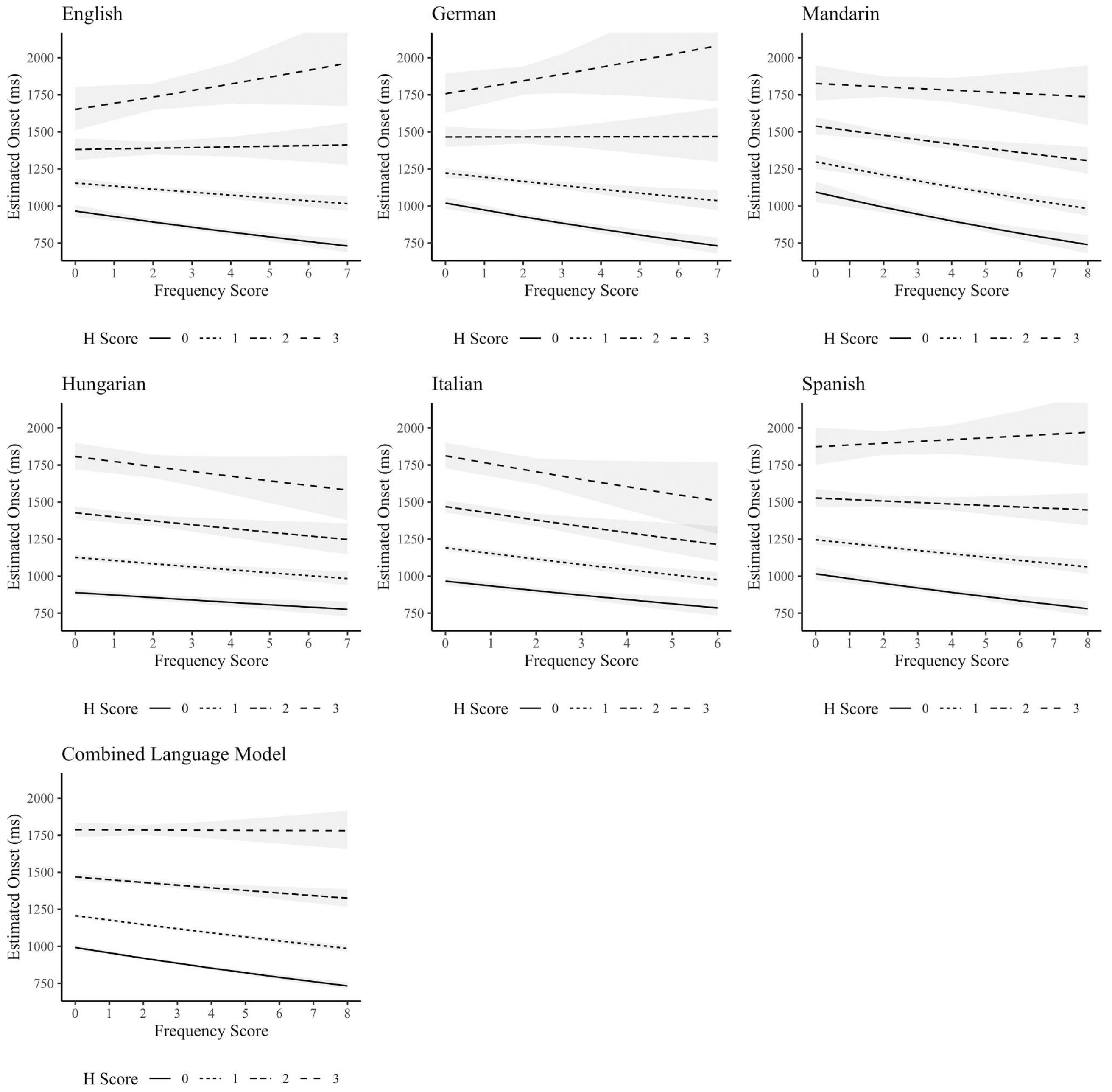

Figure 6. Codability (H score) x frequency interaction plots for the reanalysis of the Bates et al. (2003) data. RT estimates have been back-transformed from the analysis scale to the response scale (ms). Ribbons indicate 95% confidence intervals.

different stimulus set. We observed interactions between codability and frequency measures in the same language investigated in Study 1 (English) as well as in three additional languages (German, Mandarin Chinese, and Spanish). These interactions followed the same pattern as the interaction effects observed in Study 1: the frequency effect attenuated as codability decreased. We also observed a similar pattern when combining the data from multiple languages.

We did not observe a reliable under-additive interaction in all languages investigated, however: there were two languages (Hungarian and Italian) in which there was no reliable interaction. While it is possible that these cross-linguistic differences may be attributable to properties of the languages themselves or to different processing strategies in different languages, we believe that the lack of an interaction is most likely due to the frequency measures used for these languages (see discussion in Bates et al., 2003). Specifically, the Hungarian and Italian frequency measures were each derived from a corpus of approximately 500,000 words, which is much smaller than is typically used in psycholinguistic research (Füredi & Kelemen, 1989 for Hungarian; De Mauro et al., 1993 for Italian). In contrast, the frequency measures in the languages that displayed an interaction were based on corpora of approximately 2 million words or greater (Baayen et al., 1995 for English and German; Chinese Knowledge Information Processing Group, 1997 for Mandarin; Alameda & Cuetos, 1995 for Spanish). In addition, the mean frequency scores for Hungarian and Italian were lower than for the languages that displayed an interaction (Table 4). A less representative frequency measure with fewer observations in the high frequency range may make it more difficult to detect variations in the size of the frequency effect at different levels of name agreement.

Another property of the Bates et al. (2003) data set that may have influenced our ability to detect interactions between codability and frequency effects is that the data set provides mean RT per item rather than RT data for the individual elicited responses. By analysing summary statistics rather than the measures for the individual responses, we lose some of the granularity in the analysis, which could potentially obscure trends. In fact, the present analyses failed to replicate effects of word length found in other studies (e.g. D'Amico et al., 2001; Johnson et al., 1996; Székely et al., 2003; Székely et al., 2005), which supports the hypothesis that the measures used may not be sensitive enough to capture all naming RT trends.

The fact that we observed under-additive interactions between codability and frequency measures in several languages from the Bates et al. (2003) data suggests that the interactions observed in Study 1 were not merely by-products of the particular stimulus set we chose. This is what one would expect if the interaction results from an informational cascade that is a fundamental property of the language production architecture. In Study 3, we further test this hypothesis by simulating how the relationship between lexical selection and phonological encoding processes influences response time.

Study 3: simulating the source of the interaction between codability and frequency

In Study 3, we investigated whether an interaction between codability and frequency can arise as a natural consequence of an architecture that allows for cascading activation. We conducted two simulations which predicted naming time based on H score and word frequency. Our simulations focused on estimating how the relationship between lexical selection and phonological encoding influences naming RT; other processes that affect naming response time, such as conceptual access and articulation planning, are orthogonal to our primary question of how information flows between the lexical selection and phonological encoding stages of word planning. By including only lexical selection and phonological encoding in our simulations, we can get a sense of how the interplay between these processes is able to impact RT. If the relationship between these two processes on its own produces an interaction similar to those observed in Study 1 and Study 2, that provides a proof of concept that the particular architectural choice simulated could underlie the adult and child naming behaviour we observed.

We conducted both a serial simulation intended to approximate strictly sequential activation of these two levels of representation as well as a dynamic simulation intended to approximate a simple information cascade between lexical selection and phonological encoding. We analysed the RTs generated by these simulations to see whether they produced interactions between H score and frequency comparable to those observed in Studies 1 and 2, which would demonstrate that such an interaction can arise from the simulated relationship between lexical selection and phonological encoding. For the purposes of the simulations, we assume that H score influences how long it takes a speaker to select a name for articulation during lexical selection (Alario et al., 2004; Griffin, 2001) and that word frequency score influences the duration of phonological encoding (Griffin & Bock, 1998). Consequently we used H score and frequency score to estimate the durations of lexical selection and phonological encoding, respectively.

In our serial simulation, the lexical selection and phonological encoding processes were sequential, with phonological encoding taking place only after lexical selection was complete (Figure 7). In this simulation, naming RT was equal to the sum of the estimated durations of the lexical selection and phonological encoding processes. This simulation was intended to approximate a strict discrete serial activation architecture of word planning, in which activation only spreads to phonological form representations after a lexical representation has been selected for articulation.

In our dynamic simulation, phonological encoding of the name to be produced was initiated at the same time as the lexical selection process, approximating a cascading architecture of word planning in which phonological form activation begins shortly after the corresponding lexical representations are first activated (Figure 8). In the dynamic simulation, naming RT for a response was determined based on the relative estimated durations of lexical selection and phonological encoding. If the estimated duration of lexical selection was longer than the estimated time to encode the name to be produced, then the RT was equal to the duration of lexical selection (Figure 8a). This outcome simulates a situation in which

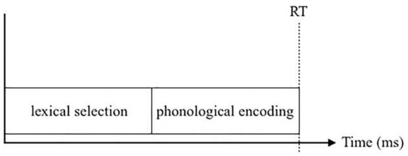

Figure 7. Schematic showing the relationship between lexical selection time and phonological encoding time on RT in the serial simulation. The dotted line represents the naming RT of a hypothetical response. The blocks labelled lexical selection and phonological encoding represent the durations of each process for that response. This simulation approximates a strict serial planning architecture in which phonological encoding does not start until after lexical selection is complete.

the phonological form of the word ultimately articulated is fully activated via cascading activation even before the speaker has decided which candidate word to produce. If the estimated duration of lexical selection was shorter than the estimated phonological encoding time, then the RT was equal to the duration of phonological encoding (Figure 8b). This outcome simulates a situation in which lexical selection is completed before the phonological form of the word to be articulated is fully activated, in which case the word cannot be not produced until encoding of its form is completed. If the estimated durations of lexical selection and phonological encoding were equal, the RT was equal to that duration (Figure 8c).

Although these simulations are simplifications of the word planning process (not intended to carefully reconstruct the complexity of real-world language production), if the simplified models can give rise to an under-additive interaction between codability and frequency, that would provide preliminary evidence that the interactions observed in Studies 1 and 2 can arise as a natural consequence of the corresponding planning architectures.

Methods

Generating hypothetical responses

The hypothetical responses in our simulations were minimally defined: the only properties attributed to each response were a word frequency score and a referent H score. To ensure that any observed RT effects in the simulations do not result from imbalances in the distribution of these properties within the data set, our hypothetical data sample was constructed in a balanced grid such that the data included samples with possible combination of frequency and H score values. Frequency score values ranged from 0.0–10.6 (the range of frequency scores of tokens in the SUBTLEX-US corpus; Brysbaert & New, 2009), and H score values ranged from 0.0–3.6 (the range of H scores in Bates et al., 2003s data set, collapsing across all languages). Both variables were incremented within our grid by 0.1, leading to a sample of 3959 hypothetical responses.

Estimating lexical decision and phonological encoding duration

We used the Bates et al. (2003) data set to approximate the marginal effects of H score and frequency score on RT, which we then used to estimate the lexical selection and phonological encoding times (respectively) for our hypothetical responses. We combined the data for all six languages in Bates et al. (2003) included in the Study 2 analysis and constructed a linear model predicting naming RT in log milliseconds. We analysed the mean RT for dominant name productions in order to maximise the accuracy of the frequency and word length measures in the data set. The model had fixed effects of H score, dominant name frequency score, and language (with a three-way interaction) as well as a fixed effect of dominant name syllable count. We applied the *Effect()* function from the {effects} package v.4.2-0 (Fox & Weisberg, 2018) to extract the marginal effects of H score and frequency score from the model. We then used the regression lines of these marginal effects to estimate the influence of H score and frequency score on log RT. The regression equation for H score was $y = 0.1950053x + 6.816174$, and the equation for frequency score was $y = -0.02744063x + 7.047661$.

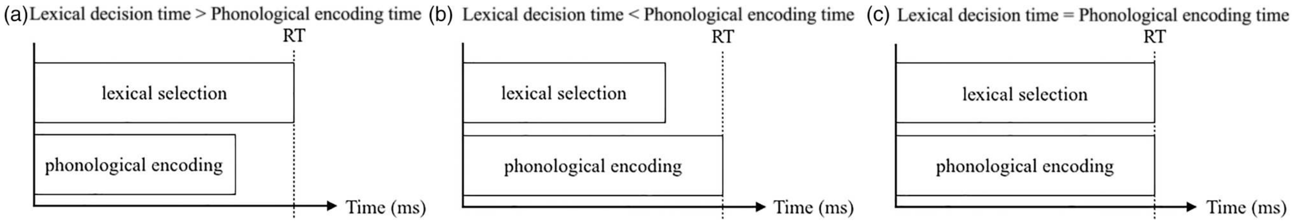

Figure 8. Schematics showing the relationship between lexical selection time and phonological encoding time on RT in the dynamic simulation. The dotted line represents the RT of the response. The blocks labelled lexical selection and phonological encoding represent the durations of each process for that response. This simulation approximates a planning architecture in which lexical selection and phonological encoding begin simultaneously.

For each hypothetical response, we entered its H score and frequency score into the corresponding linear equation and exponentiated the resulting values to obtain time estimates in milliseconds. We used these time outputs as our estimates of the duration of the lexical selection and phonological encoding processes. For example, a hypothetical response with an H score of 1.0 and a frequency score of 1.0 would have an estimated lexical selection duration of *exp (0.1950053*1.0 + 6.816174)*, or approximately 1109ms, and an estimated phonological encoding duration of *exp(−0.02744063*1.0 + 7.047661)*, or approximately 1119 ms.

Calculating naming RT

We used the lexical selection and phonological encoding durations for each hypothetical response to determine its naming RT in our simulations. The same set of hypothetical responses was used in both simulations. In these simplified simulations, a response was considered "produced" (i.e. named) as soon as (i) its lexical duration had elapsed (i.e. simulating that the response name had been selected at the lexical level) and (ii) its phonological encoding duration had elapsed (i.e. simulating that the response's phonological form fully retrieved and encoded).

In the serial simulation, the RT of a response was equal to the sum of its lexical selection duration and its estimated phonological encoding duration. For example, if a response had a lexical selection duration of 1200 ms and a phonological encoding duration of 1100 ms, its RT would be 2300 ms in the serial simulation. In the dynamic simulation, if a response's lexical selection duration was longer than its estimated phonological encoding duration, its RT was equal to its lexical selection duration. For the above example, the response's RT would thus be 1200 ms in the dynamic simulation. If a response's estimated lexical selection duration was shorter than its estimated phonological encoding duration, the RT was equal to the phonological encoding duration. For example, for a response with a lexical selection duration of 1200 ms and a phonological encoding duration of 1250 ms, the RT would be 1250 ms. If a response's lexical selection and phonological encoding durations were equal, the RT was equal to that duration.

Analysis

We analysed the RTs produced by the two simulations for the hypothetical responses using linear regression models. To parallel the RT analyses in Studies 1 and 2, we analysed log-transformed RTs (log milliseconds). We constructed separate regression models for each simulation. These models had fixed effects of H score and frequency score with an interaction between them.

Results

The RTs were on average longer in the serial simulation ($M = 2323$ms, SD = 287 ms, range = 1772–2991 ms) than in the dynamic simulation ($M = 1335$ ms, SD = 262 ms, range = 913–1841 ms). This is an expected by-product of the way that the RTs were calculated.

Figure 9 shows the relationships between H score, frequency score, and RT predicted by the analysis models for the serial and dynamic simulations. The shape of this relationship was very different in the two simulations. The interaction plot for the serial simulation shows no obvious change to the slope of the frequency effect as H score increases. In contrast, the shape of the interaction plot for the dynamic simulation closely resembles the interactions between H score and frequency observed in Studies 1 and 2; as H score increases, the negative slope of the frequency effect decreases.

The main effects of H score and frequency score were significant for both the serial simulation (H score: $\beta = 0.10$, $t(3955) = 728.79$, $p < 0.0001$; frequency score: $\beta = -0.01$, $t(3955) = -293.30$, $p < 0.0001$) as well as the dynamic simulation (H score: $\beta = 0.15$, $t(3955) = 215.33$, $p < 0.0001$; frequency score: -0.01, $t(3955) = -50.81$, $p < 0.0001$). RTs were shorter for responses with lower H scores (higher codability) and for responses with higher frequency scores.

The interaction between H score and frequency was significant in the dynamic simulation ($\beta = 4.8\text{e-}03$, $t(3955) = 41.49$, $p < 0.0001$), where the interaction effect plot shows an attenuation of the frequency effect as H score increases (Figure 9). Surprisingly, this interaction was also significant for the serial simulation ($\beta = 1.3\text{e-}03$, $t(3955) = 56.48$, $p < 0.0001$), where there is no obvious change in the slope of the frequency effect by H score value in the interaction plot (Figure 9). Further exploration of the data suggests that the interaction in the serial simulation (but not the dynamic simulation) is a by-product of log transforming the dependent variable in the analysis. Repeating our analyses using untransformed RTs, the interaction between H score and frequency score is no longer significant in the serial simulation ($\beta = 0.00$, $t(3955) = 0.00$, $p = 1$), even though the relationship between H score, frequency score, and RT appears almost identical to that in Figure 9 (see Supplementary Materials). By contrast, the pattern of results in the dynamic simulation remains the same when analysing untransformed RTs (see Supplementary Materials). Importantly, as in the dynamic simulation, we continue to see evidence of under-

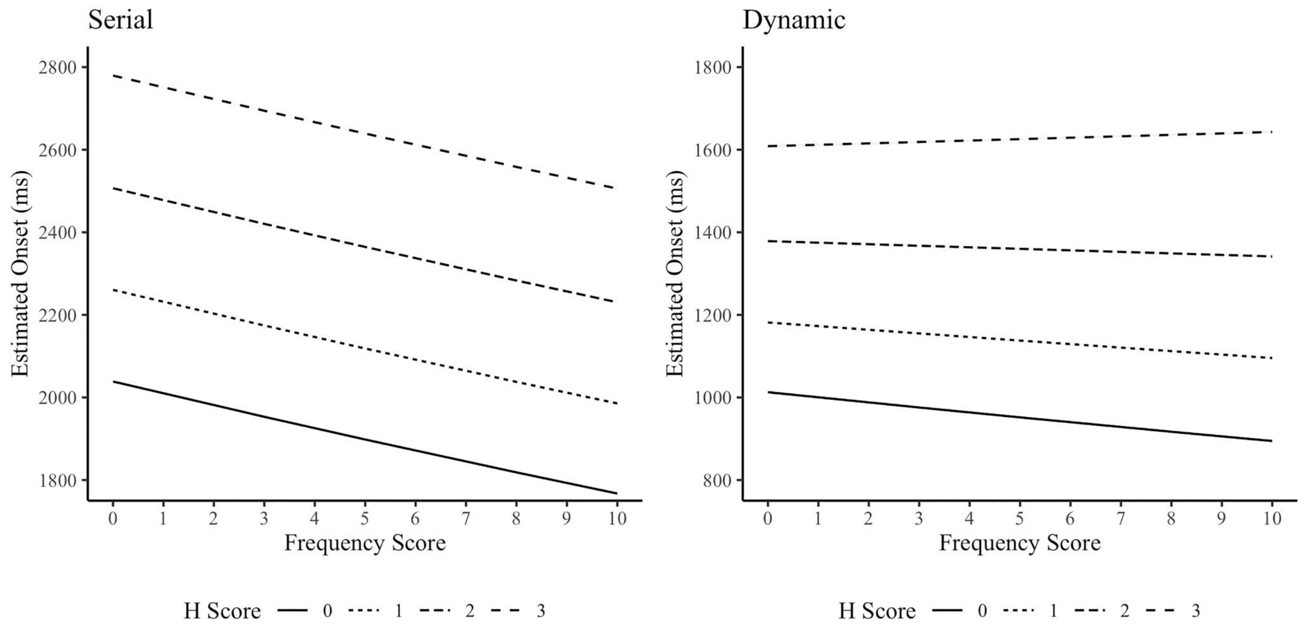

Figure 9. Interactions between H score and frequency score in the Study 3 simulations. RT estimates have been back-transformed from the analysis scale to the response scale (ms). Ribbons indicate 95% confidence intervals.

additive interactions between H score and frequency score measures in the Study 1 data when using untransformed RTs for analysis; we observe interactions in both the adult and child data (though the interaction only reaches conventional levels of significance in the adult data when using a simplified random effects structure; see Supplementary Materials).

To further explore whether the serial simulation produces an under-additive interaction between codability and frequency measures similar to those observed in Study 1, we performed a categorical analysis of the simulated data in the style of the Study 1 categorical analysis (see Supplementary Materials). In this analysis, there was no significant interaction between codability and frequency category in the serial simulation ($\beta = -3.26\text{e-}03$, $t(3955) = -1.12$, $p = 0.26$). In the dynamic simulation, there was a significant under-additive interaction similar to those observed in Study 1, with an attenuated frequency effect when codability is low ($\beta = -0.03$, $t(3955) = -5.96$, $p < 0.0001$). This pattern of results does not change when using untransformed RT as the dependent variable.

Study 3 discussion

In Study 3, we compared the interaction between referent H score and word frequency score in two simulations with different relationships between lexical selection and phonological encoding. In the serial simulation, phonological encoding began only after lexical selection was complete. In the dynamic simulation, phonological encoding began at the same time as lexical selection, simulating a simple informational cascade between representations at the two processing levels.

The dynamic simulation produced response times that exhibited an attenuation of the frequency effect as H score increased (Figure 9). This interaction between codability and frequency had a similar shape to those observed in Studies 1 and 2. The response times produced by the serial simulation, on the other hand, did not show the same under-additivity between codability and frequency (Figure 9).

Surprisingly, in our primary analyses, the interaction between H score and frequency score was significant for both simulations. However, we recommend that the interaction in the serial simulation analysis be interpreted with caution. Unlike in the dynamic simulation, there is no comparable interaction for the serial simulation when analysing the simulated RTs in a categorical fashion. Moreover, the interaction in the continuous analysis of the serial simulation appears to be a by-product of analysing log-transformed RT. The interaction in the serial simulation does not persist in an analysis of untransformed response time, even though the relationship between H score, frequency score, and RT appears virtually identical in the modelled interaction (see Supplementary Materials for plot). Critically, the interaction *does* persist when analysing untransformed RTs in the dynamic simulation as well as the Study 1 data. Nevertheless, even if one assumes that the interaction between H score and frequency score is reliable in the serial simulation, the shape of this interaction (Figure

9) is so drastically different from those observed in Studies 1 and 2 that it would be unreasonable to assume that they result from the same underlying relationship between H score, frequency score, and RT.

The fact that a clear and reliable under-additive interaction between codability and frequency is produced by a simulation of word planning in which phonological encoding of the produced word begins before lexical selection has been completed (but not by a simulation when the two processes occur sequentially) supports the hypothesis that the interaction effects observed in Studies 1 and 2 can arise as a natural consequence of cascading activation between lexical and phonological forms during word planning.

It is important to note, however, that the simulations in Study 3 have limitations to their psychological plausibility. We identify several of these in Table 6 along with the potential consequences they have for our interpretation of the simulations. We conclude that these simplifications are unlikely to account for the differences between the two simulations and are unlikely to prevent the simulations from capturing the relationship between codability and phonological frequency effects. We draw the reader's attention to the final two of these limitations, which, on the face of it, seem most relevant to our theoretical interpretation of the interaction between codability and frequency.

First, the simulations do not model any competition effects between forms at the phonological encoding level (note that competition at the lexical selection level is modelled by the H score effect). We know that phonological competition does occur during picture naming: reaction times are longer for names that are in more dense phonological neighbourhoods (e.g. Sadat et al., 2014; Zhang et al., 2020). Nevertheless, it seems unlikely that phonological competition would fundamentally change the behaviour of the simulations or offer an alternative explanation for the critical interaction.

On the one hand, adding phonological competition to the serial simulation would not, in and of itself, produce an under-additive interaction between codability and frequency, since the effects of codability would necessarily occur before and separate from phonological encoding. In a strict serial architecture, phonological competition would come from phonological neighbours of the target word to be articulated. We would not expect an under-additive interaction unless the names applied to referents with low codability reliably have larger phonological neighbourhood densities than those applied to referents with higher codability, leading to differential slowing at different levels of name agreement. This does not appear to be the case: investigating the relationship between H score and dominant name phonological neighbour count in Bates et al.'s (2003) English, German, and Spanish data for single word dominant names, we did not observe such a trend (phonological neighbourhood size was derived from CLEARPOND; Marian et al., 2012; CLEARPOND data was not available for the other languages in the data set). In the Spanish data, H scores were not significantly correlated with dominant name phonological neighbourhood count (Pearson's $r = 0.03$, $t(501) = 0.77$, $p = 0.44$), and there was a significant negative correlation in the English (Pearson's $r = -0.12$, $t(518) = -2.66$, $p < 0.01$) and German (Pearson's $r = -0.08$, $t(518) = -2.02$, $p = 0.04$) data, meaning that phonological neighbourhood size was greater for items with higher codability (more name agreement). A negative relationship between phonological neighbourhood and name agreement predicts attenuation of the frequency effect at *higher* levels of codability rather than lower ones, producing the opposite pattern of the codability and frequency interaction we observed in Studies 1 and 2.

On the other hand, adding phonological competition to the dynamic simulation is unlikely to eliminate the under-additive interaction. If multiple phonological forms receive spreading activation from an informational cascade, one might expect increased competition between these activated forms, even after only one lexical representation is selected for articulation. In fact, in Study 1 participants produced more false starts in the low codability conditions than the high codability conditions (see Supplementary Materials), which could reflect an influence of activated word forms other than the one the speaker intended to produce. However, in the dynamic simulation, adding phonological competition is more likely to accentuate the under-additive interaction than to erase it. Specifically, in the dynamic simulation, phonological competition would likely be greater when name agreement is low, allowing for phonological competition from unselected candidates (and their phonological neighbours). This would lead to slower response times for low codability pictures, particularly when the preferred label was low frequency (relative to the competitors). This slowing may further attenuate any potential processing boost the target has from increased form frequency at higher H score levels, resulting in a greater under-additive effect. Consequently, we do not believe adding phonological competition to either the serial or dynamic simulation would change the pattern of observed results.

A second limitation of the simulations, and the one that seems most relevant to the critical interaction, is that they do not account for frequency effects at the level of lexical selection. As mentioned in the *Introduction*, there is evidence that word frequency may influence lexical selection processes in addition to the selection of phonological form (e.g. Finocchiaro & Caramazza, 2006; Jescheniak & Levelt, 1994; Johnson et al., 1996; Kittredge et al., 2008; Strijkers et al., 2010). By attributing frequency effects solely to phonological encoding, the simulations may overestimate the size of the phonological frequency effect, though as mentioned in Table 6, we are not concerned about capturing the precise magnitude of this effect. More crucially, if both frequency and name agreement influence the process of lexical selection, that creates an opportunity for the two factors to interact at that level. While there is no particular reason to assume that the factors would interact, or that this interaction would be under-additive, this does open the possibility that a serial simulation with these features might produce the critical interaction. We discuss this possibility further in the *General discussion*.

Despite these simplifications, the simulations provide a rough sketch of how referent name agreement and word frequency may interact in word planning architectures with different relationships between lexical selection and phonological encoding. The simplicity of the simulations allows us to see how the relationship between these two processes can influence naming RT in the language production architecture under idealised conditions. The simulations show that an under-additive interaction between codability and frequency, similar to that observed in Studies 1 and 2, arises naturally from a planning architecture in which phonological planning begins before lexical selection has been completed, approximating an informational cascade between the two processing levels. Coupled with the prior evidence for cascaded processing in adults (see *Introduction*) and the qualitatively similar naming behaviour between adults and five-year-olds in Study 1, we believe that these results support the hypothesis that cascading activation is present in the five-year-old language production system.

General discussion

In the present study, we analysed naming times in five-year-old children and adults to understand how information flows between levels of representation in both the developing and mature language production systems. We asked whether the informational cascades present in the adult production architecture are a fundamental property of the language system present early in life or whether they only emerge later with experience. We addressed these questions in three studies.

Study 1 assessed the influence of codability (name agreement) and frequency manipulations on picture naming response time by adults and five-year-old children. By investigating these two factors, which have been argued to influence the processes of lexical selection (Alario et al., 2004; Griffin, 2001; inter alia) and phonological encoding (see Griffin & Bock, 1998), respectively, we can assess how these two processes relate to each other. Replicating prior results, we found that naming RTs were affected by both manipulations in both populations (Study 1, Q1). We observed slower naming times for items with low codability and for low frequency names (see also Bates et al., 2003; Butterfield & Butterfield, 1977; D'Amico et al., 2001; Jescheniak & Levelt, 1994; Johnson, 1992; Johnson & Clark, 1988; Lachman, 1973; Lachman et al., 1974; Lachman & Lachman, 1980; Oldfield & Wingfield, 1965; Paivio et al., 1989; inter alia). Fitting ex-Gaussian distributions to the adult and child RT data (Study 1, Q2), we found that the codability and frequency manipulations engendered different effects on the RT distributions, supporting the hypothesis that they influence different processes. The RT distribution effects of the manipulations were consistent with their interpretations of influencing the processes of lexical selection and phonological encoding, respectively. These effects were qualitatively similar in both age groups, suggesting that there are similar underlying production planning processes at play in both the developing and mature language systems.

Critically, we also observed significant under-additive interactions between codability and frequency effects in both the adults and the five-year-olds: the size of the frequency effect was reduced when codability was low. To our knowledge, such an interaction has not been previously reported for either population. In Study 2, we demonstrated that this interaction is reliable across experiments and languages, documenting the same pattern of effects in Bates et al.'s (2003) English, German, Spanish, and Mandarin Chinese naming data. This under-additive interaction is predicted by a cascading model of word planning in which activation spreads from activated lexical representations to their phonological forms even before a lexical item has been selected for articulation (more below).

In Study 3, we confirmed our hypothesis about the source of the interaction by simulating the codability effect on lexical selection and the frequency effect on phonological encoding in different word planning architectures. We conducted two simulations: one in which lexical selection and phonological encoding processes occur simultaneously (approximating a simple informational cascade) and another in which they occur sequentially (approximating a discrete, serial planning process). The simulations showed that an under-additive interaction between codability and frequency measures

Table 6. Limitations of the Study 3 simulations.

Limitation	Potential consequences
The simulations simplify the word production process to two steps (lexical selection, phonological encoding) and do not take into account other processes involved in naming (e.g. conceptual processing, articulatory planning).	If these other processes are serial and separate, then our simulations capture the naming process with reasonable accuracy (just add constants for each process and some noise). If these processes are also subject to an informational cascade, this may provide other possible loci for interactions not accounted for in the simulations.
The simulations leave out other factors that influence picture naming RT (e.g. AoA, word length, by-speaker variation).	If these factors are orthogonal to the effects of interest, then this variation would simply be an additional source of noise in the response times and would not affect the presence/absence of an interaction between codability and frequency. If these factors account for some of the same variability as codability and/or frequency measures, they may be other possible contributors to interactions that are not accounted for in the simulations.
The simulations simplify the estimation of codability and frequency effects on lexical selection and phonological encoding as linear equations. These equations are unlikely to capture the precise magnitude of the effects for all individuals, circumstances, and languages.	For the goals of the present analysis, capturing the exact magnitude of the slowdowns in lexical selection and phonological encoding processes caused by codability and frequency is less important than capturing the way they interact. In additional simulations, we found that the overall pattern of results (no interactions for serial simulations, under-additive interactions for dynamic ones) is robust to different estimations of lexical selection and phonological encoding times. This suggests that the precise magnitude of the estimated codability and frequency effects in the simulations does not in and of itself produce the presence or absence of an interaction.
The simulations do not account for processing limitations such as processing load, working memory capacity, retrieval errors, or confusion.	These sources of noise in response times may be more likely to influence naming when codability is low (i.e. when there are more active lexical representations). If this is the case, these factors should make the codability effect more pronounced and may slow other processes like phonological retrieval when there is low name agreement, potentially contributing to an *over-additive* interaction between lexical selection and phonological encoding processes. If such processes are at play in the real world and retrieval is otherwise serial, then we'd expect no interaction or an over-additive interaction. This is ruled out by the data. If these processes exist in the context of cascading architecture, then our simulations (which do not include these effects) would overestimate the under-additivity of the observed interaction in the dynamic simulation, resulting in a more pronounced fan shape. Retrieval errors that slow RTs may be less likely to occur for more frequent forms, which should make the frequency effect more pronounced. How this affects the interaction is dependent upon the locus of this frequency effect.
The simulations assume clean and instantaneous transmission of information between lexical selection and phonological encoding processes. The simulations do not assume an activation threshold for lexical items that must be reached before activation spreads to phonological form.	Non-instantaneous information transfer is an additional source of noise in the response times that is unlikely to cause or inhibit an interaction between codability and frequency. An activation threshold for information spread may increase the length of lexical selection by requiring that additional processing must occur before phonological encoding can start. This lengthening may be more likely in cases of low codability when there are more candidate names receiving activation. Lengthening lexical selection in cases of low codability is unlikely to produce an interaction in a serial framework. In a dynamic framework, this lengthening may lead to an even greater under-additive interaction.
The simulations do not model the potential effect of phonological competition on the speed of phonological encoding (e.g. in a resource-limited parallel model) or the potential role of activation feedback from phonological forms to their lexical representations.	In the dynamic simulation, adding phonological competition could potentially accentuate an under-additive interaction between codability and frequency, as competition would result in increased slowing in cases of low codability (when the forms of many lexical items become activated) that may attenuate any potential processing boost the target has from form frequency. In the serial simulation, if effects of codability occur before and separate from phonological encoding, adding phonological competition would not produce an under-additive interaction between codability and frequency on its own. We would not expect phonological competition to attenuate the size of the frequency effect for low name agreement (leading to an under-additive interaction) unless phonological neighbourhood densities are larger for names applied to referents with lower codability.
The simulations do not account for an effect of frequency on lexical selection.	The simulation likely overestimates the size of the phonological frequency effect. If frequency influences lexical selection in addition to codability, that provides the opportunity for the two factors to interact at that level, potentially producing an interaction within a serial architecture that is unaccounted for in our simulation. In the dynamic simulation, distributing the frequency effect across the lexical selection and phonological encoding processes should not eliminate the under-additive interaction between codability and the phonological frequency effect, though it may make the interaction less pronounced if the phonological frequency effect is smaller.

arises naturally in the first simulation but not the second. While this cascading model is not the only possible explanation for this interaction, we believe that given the prior evidence for cascaded processing in adults it is the most parsimonious explanation (e.g. Costa et al., 2000; Cutting & Ferreira, 1999; Jescheniak & Schriefers, 1998; Morsella & Miozzo, 2002; Peterson & Savoy, 1998; Rapp & Goldrick, 2000; Starreveld & La Heij, 1995; among many others). Thus, the presence of the same interaction in five-year-olds (Study 1) provides evidence that cascaded processing is already robustly present by this age, as we would expect it to be if this cascade is an inherent consequence of an incremental language production system.

In the remainder of the *General discussion*, we discuss potential interpretations of the interaction between codability and frequency effects (including alternative interpretations that do not assume cascaded processing), what our findings suggest for the development of cascading activation in the language production system, and questions for future research.

Interpreting the codability and frequency interaction

The presence of an interaction between image codability and response frequency in picture naming is intriguing, as it requires that the underlying mechanisms affected by the two factors be able to influence each other. This is difficult to reconcile with strictly sequential models of language production because codability and frequency effects have generally been assumed to affect different processes within word planning (lexical selection and phonological encoding, respectively). We propose that the observed interaction reflects an incremental informational cascade between lexical selection and phonological encoding processes.

In a cascading framework, activation spreads to the phonological forms of the lexical candidates activated during selection. When codability is low and lexical selection time is lengthened, the phonological forms of the name alternatives (including the name ultimately selected for articulation) have more time to receive a cascading activation boost before selection of a lexical candidate. Given that the frequency of a word form influences the time it takes to access the phonological form, the size of the frequency effect should be diminished when the form of the selected name is already partially activated before phonological encoding, compared to cases when the word form receives less cascading activation and starts phonological encoding from closer to its base activation level. Viewed from a different perspective, when lexical selection is slowed due to low codability, word frequency's RT influence on phonological encoding will become less pronounced, as relatively faster or slower access of phonological forms will be in part obscured by the slowing at the lexical selection level. Thus, a cascading activation architecture predicts the under-additive interactions we observed.[9]

A strict serial architecture, on the other hand, does not predict that the factors influencing lexical selection and phonological encoding will interact. In such an architecture, activation spreads to the selected word's phonological form only once competition between different lexical candidates is resolved during lexical selection. In this framework, codability should influence the speed of lexical selection, with longer selection times when there are more active lexical candidates (i.e. codability is low), but it should not influence the subsequent phonological encoding stage once a name candidate has already been selected for utterance. Thus, under a strict serial model of word planning, the two effects should be additive (see Sternberg, 1969, 2001 for discussion).

Nevertheless, the presence of such an interaction, interpreted in isolation, does not provide conclusive evidence for cascading activation. There are (at least) two other ways that such an interaction could arise. The first way weakens the central assumption of the serial activation hypothesis by allowing multiple lexical nodes to be selected but only when multiple names could apply to an image (e.g. Levelt et al., 1999), as is the case for low codability items. In this scenario, the phonological forms for multiple candidate names may be activated for low codability items, which could lead to increased competition between forms at the phonological level. What effect this would have on the frequency effect would depend on the frequency distribution of the alternative labels and the degree to which frequency plays a role in resolving the competition at this lower level. If competition decreases the frequency effect (by introducing other factors that limit phonological encoding and produce overall slowing), this model *might* result in an under-additive interaction and thus capture the present data pattern.

Second, the interaction could result from an interplay of codability and frequency effects during lexical selection. As previously mentioned, there is evidence that frequency may influence lexical selection in addition to phonological encoding (e.g. Finocchiaro & Caramazza, 2006; Jescheniak & Levelt, 1994; Johnson et al., 1996; Kittredge et al., 2008; Strijkers et al., 2010). For example, a frequency effect at the lexical selection level may enhance the base activation of more frequent lexical candidates (Strijkers et al., 2010), which could allow for faster activation and competition resolution in favour of these frequent candidates. Such an effect has the potential to speed up selection (and consequently production) of high frequency lexical representations, particularly in cases of high codability when there is little competition from alternative lexical candidates. Highly codable items with high frequency names should thus receive frequency-related activation boosts during both the lexical selection and phonological encoding processes, speeding their RTs compared to their low frequency counterparts. It is possible that a lexically-based frequency effect would have a smaller effect when codability is low and there is greater competition from alternative lexical candidates, leading to an under-additive interaction. To support this alternative account over a cascading hypothesis, one must determine whether the size of a lexical selection-based frequency effect is large enough to produce an interaction effect on its own (frequency effects on lexical selection have been argued to be smaller and less reliable than the frequency effect on phonological encoding; Griffin & Bock, 1998) and whether the size of such an effect is indeed diminished when there are multiple lexical candidates.

In sum, while an informational cascade between lexical selection and phonological encoding is one possible explanation for the interaction we observed, this explanation is not the only way such an interaction could arise (particularly given the complexity of attributing effects of manipulated variables to specific levels of processing). Nevertheless, this explanation relies on the fewest unproven auxiliary assumptions and is consistent with the other evidence for cascading processing in adults (e.g. Costa et al., 2000; Cutting & Ferreira, 1999; Jescheniak & Schriefers, 1998; Morsella & Miozzo, 2002; Peterson & Savoy, 1998; Rapp & Goldrick, 2000; Starreveld & La Heij, 1995).

Cascading activation in the developing language production system

Our findings in adults provide convergent evidence for a theory of mature language production that is well supported by prior studies of speech errors and interference paradigms – a theory in which there is cascading activation between lexical selection and phonological encoding. Our findings for children demonstrate that this same theory can explain the pattern of picture naming times in five-year-olds. Specifically, in young children, as in adults, there is an under-additive interaction between the codability and frequency effects, which is most readily explained by the continuous spread of activation from lexical selection to phonological encoding. This is important because prior studies directly addressing this question have been limited to children seven years of age and older (e.g. Jescheniak et al., 2006; Poarch & van Hell, 2012; Sylvia, 2017). The fact that cascaded processing is present this early in development suggests that it is a fundamental property of the language system as opposed to a property that emerges only with experience or adult-like cognitive processing.

Critically, however, the present study does not resolve the question of when and how cascaded processing develops in young children; it merely places new constraints on the answer. Children begin mapping words to meanings as young as six months of age (Bergelson & Aslin, 2017; Bergelson & Swingley, 2012; Bergelson & Swingley, 2013; Bergelson & Swingley, 2015; Tincoff & Juscyzk, 1999; Tincoff & Juscyzk, 2011). Their interpretation of words appears to improve substantially by 13–14 months (Bergelson & Swingley, 2012; Bergelson & Swingley, 2013; Bergelson & Swingley, 2015). In the second year of life, there are further improvements in the speed of lexical processing (Fernald et al., 1998; Fernald et al., 2006; Zangl et al., 2005) and changes in the robustness of phonological representations (Mills et al., 2004; Werker et al., 2002). Thus, comprehension studies suggest that the most substantial changes in lexical representations may be occurring between roughly 12 and 24 months of age. Does cascading processing in lexical production emerge with these early changes in language comprehension? Or does it rely on later changes in the lexical network, like those associated with lexical prediction (Mani & Huettig, 2012)? The present study provides strong motivation for pursuing these questions in even younger children.

In addition, our findings bear on the question of how the effects of cascading activation change during the middle childhood and adolescence. As we noted in the introduction, Jescheniak et al. (2006) observed preliminary evidence for a developmental change in cascading activation. They found evidence of semantically-mediated phonological activation (a predicted consequence of a cascaded processing architecture) in seven and eight-year-old children but not in nine and ten-year-old children or adults, which could suggest that the informational cascade between lexical representations and phonological forms is more robust in younger children (Jescheniak et al., 2006). Similarly, visual inspection of the interaction effect plots from our naming experiment (Figures 3 and 4) suggests that the interaction between codability and frequency may be more pronounced for five-year-olds than for adults. This observation is supported in the categorical analysis by the relative magnitudes of the beta coefficients for the interaction in the adult ($\beta = -0.03$) and child ($\beta = -0.06$) data. Moreover, in an exploratory categorical analysis, there was a three-way interaction between codability, frequency, and population ($\beta = 0.02$, $t(160) = 3.66$, $p < 0.001$) (though this contrast is not obviously reliable using continuous measures; see Supplementary Materials). If we assume that the interaction between codability and frequency reflects an informational cascade, this pattern would also suggest a developmental change in the strength of these effects.

A stronger cascade earlier in development could result from weaker inhibition of cascaded activation in younger children, resulting from a general deficiency in inhibitory capacity compared to older children and adults (see discussion in the *Introduction*). Under this hypothesis, evidence of a cascade should become less pronounced as speakers become better able to suppress the activation of non-target forms. In fact, the child participants in Study 1 produced a greater proportion of false starts and renaming errors in the low codability conditions (1.6% of responses) than adults (0.6%), which may reflect such weaker inhibition of alternative name candidates (see Supplementary Materials for error distributions). Alternatively, a stronger cascade in younger children could result from slower overall processing, allowing more activation to cascade to non-target forms (see *Introduction*); as production processes become more efficient and faster over development, non-target forms may receive less activation from the informational cascade, making their activation more difficult to detect.

Further research is necessary to identify the developmental trajectory of cascaded processing, including assessing whether it is possible to find evidence for informational cascades in even younger children and investigating possible changes in the strength of the cascade over time. Tracking the size of the codability and frequency interaction in naming RTs across different age groups has the potential to inform whether the strength of informational cascades changes over the course of development. One of the advantages of the picture naming paradigm we used in our study is that it is simpler than the priming and interference tasks typically used to look for evidence of cascaded processing in adults. The straightforward nature of the picture naming task makes it easy to run with child populations, providing a potential avenue to explore when cascaded processing first arises.

In addition to investigating interaction effects between codability and frequency, researchers could use this paradigm to look for other signatures of phonological form activation of multiple lexical candidates. For example, researchers could investigate children's production of false start and renaming errors, which may be more likely to occur when activation cascades to candidates other than the ultimately produced target name. Future work could manipulate additional properties of the images to be named that may influence the level of competition experienced in cases of an informational cascade, such as the relative frequencies and/or the phonological neighbourhood densities of their candidate names. Tracking speech errors longitudinally may provide insight into the strength of the informational cascade and/or inhibition of such cascaded activation over development.

Nevertheless, the interaction we observed and the presence of speech errors, while consistent with an informational cascade, do not prove the existence thereof. A remaining question for future research is whether it is possible to obtain more direct evidence for informational cascades in children under seven years of age – for instance, by finding evidence of semantically-mediated phonological activation. Ideally, future research will additionally explore a wider range of data sets to assess not only when this property arises but also whether there are cross-linguistic differences in development or in the end state of the cascaded word planning architecture.

Conclusion

In the present study, we address the question of how early cascading activation emerges in the language production system. While prior work has shown evidence of informational cascades in children seven years of age and older (e.g. Jescheniak et al., 2006; Poarch & van Hell, 2012; Sylvia, 2017), our study provides preliminary evidence that a cascading architecture is already in place by at least by five years of age. In a picture naming experiment, we observed qualitatively similar response time effects in both adults and five-year-old children, suggesting that similar underlying word planning processes are at play in both the developing and adult language systems. Critically, we observed under-additive interactions between image codability and name frequency effects in both populations. This interaction generalises across experiments and languages and arises naturally from a word planning architecture in which activation cascades between lexical and phonological representations. Our study thus provides evidence for early cascaded processing, supporting the hypothesis that it is a fundamental property of the language system, rather than a capability that emerges gradually with experience.

Notes

1. Reconciling the mixed error effect with a serial model of lexical planning (e.g. Levelt et al., 1991) requires the

assumption of a post-encoding editor (Baars et al., 1975; Butterworth, 1981; Kempen & Huijbers, 1983; Levelt, 1989).

2. It is important to note, however, that codability effects, while commonly attributed to co-activation at the lexical level, may not exclusively reflect an influence on lexical decision; name agreement may also influence processes prior to lexical decision such as conceptual access.
3. One exception we have found is an adult sentence production study by Spieler and Griffin (2006). Their experiment elicited sentences in the form *The A and the B is above the C*. The researchers manipulated the frequency (high, low) and codability (high, medium) of critical items that appeared in either the *B* or *C* position (the item in *A* always had high codability). They observed an interaction between the frequency and codability of the critical items on the latency between the onset of *A* and the onset of the critical item. This interaction is not in the direction we observe, however: they observed an over-additive effect of frequency for medium codable items compared to highly codable items (latencies were especially slow for low frequency, medium codable items). While not extensively discussed in Spieler and Griffin (2006), they similarly attribute such an effect to cascading activation (see discussion in Griffin & Bock, 1998 about how increased constraint in word choice may attenuate the influence of frequency).
4. The power analysis was performed using the random effects structure specified in our preregistration.
5. The adult preregistration lists the use of an adjective as a potential example of prenominal verbalisation. Given that some of the stimulus images elicited responses with adjectives that modified the head noun such that the response could potentially be considered a single lexical item (e.g. a compound noun) with a different meaning than the head noun (e.g. school bus, steering wheel, candy bar), we decided to allow adjectives in the responses.
6. We altered the model structure specified in the preregistration to (i) make our model estimates more conservative and (ii) aid in model convergence. We added a random intercept for item as well as an interaction term in the random slope by subject, and we omitted the random effect of trial (an investigation of participant slopes for trial suggested minimal variation in trial slope by participant).
7. The large τ effect in the child data is unlikely to be solely attributable to the imbalance in target name frequency between the child codability conditions (Table 3). The adult data suggest that an effect of frequency should influence τ and μ roughly evenly, meaning that the lower frequency of the low child codability condition should not have disproportionately inflated the τ effect above the μ effect.
8. Bates et al. (2003) also elicited data from native speakers of Bulgarian, however we decided to exclude this language from our analysis, as the frequency measure provided was in the form of subjective ratings instead of derived from corpus counts. It is not clear that we should expect a frequency measure on an ordinal scale to show the same interaction with H score as the frequency score measure used in Study 1.
9. As pointed out by a reviewer, under some assumptions, cascading activation may in fact predict an over-additive interaction. In particular, over-additivity could arise if (i) phonological encoding begins as soon as lexical candidates are active, (ii) phonemes linked to different lexical candidates compete with each other (but phonemes activated by the same candidate do not), (iii) phonological competition depends on frequency (high frequency competitors result in greater competition), and (iv) in cases of low name agreement, the frequency advantage for the produced name relative to alternatives is greater for high frequency words than low frequency words. Under these assumptions, competition during phonological encoding would have a minimal effect on high frequency, low codability items and the largest effect on low frequency, low codability items, thereby resulting in an over-additive interaction. Our data are inconsistent with this hypothesis.

Acknowledgements

Thank you to Parker Robbins and Benazir Neree for their assistance with data collection and processing as well as to Alfonso Caramazza and Joshua Cetron for sharing their thoughts on the project and analyses. We are additionally grateful to the anonymous reviewers of this article for their helpful comments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by an internal grant from Harvard University, FAS Psychology.

Data availability statement

Data and Supplementary Materials are available from https://osf.io/myrtg/.

References

Abdel Rahman, R., & Melinger, A. (2008). Enhanced phonological facilitation and traces of concurrent word form activation in speech production: An object-naming study with multiple distractors. *Quarterly Journal of Experimental Psychology*, *61*(9), 1410–1440. https://doi.org/10.1080/17470210701560724

Alameda, J. R., & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano* [Frequency dictionary for lexical items in Castilian Spanish]. Servicio de Publicaciones de la Universidad de Oviedo.

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*(1), 140–155. https://doi.org/10.3758/BF03195559

Anderson, J. D. (2008). Age of acquisition and repetition priming effects on picture naming of children who do and do not stutter. *Journal of Fluency Disorders*, *33*(2), 135–155. https://doi.org/10.1016/j.jfludis.2008.04.001

Andrews, S., & Heathcote, A. (2001). Distinguishing common and task-specific processes in word identification: A matter of some moment? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(2), 514–544. https://doi.org/10.1037/0278-7393.27.2.514

Baars, B. J., Motley, M. T., & MacKay, D. G. (1975). Output editing for lexical status in artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, *14*(4), 382–391. https://doi.org/10.1016/S0022-5371(75)80017-X

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX Lexical Database* (Release 2) [CD-ROM]. Linguistic Data Consortium.

Balatsou, E., Fischer-Baum, S., & Oppenheim, G. M. (2022). The psychological reality of picture name agreement. *Cognition*, *218*, Article 104947. https://doi.org/10.1016/j.cognition.2021.104947

Balota, D. A., & Spieler, D. H. (1999). Word frequency, repetition, and lexicality effects in word recognition tasks: Beyond measures of central tendency. *Journal of Experimental Psychology: General*, *128*(1), 32–55. https://doi.org/10.1037/0096-3445.128.1.32

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language*, *59*(4), 495–523. https://doi.org/10.1016/j.jml.2007.10.004

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., Herron, D., Lu, C. C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., & Hsu, J. (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, *10*(2), 344–380. https://doi.org/10.3758/BF03196494

Bergelson, E., & Aslin, R. (2017). Nature and origins of the lexicon in 6-mo-olds. *Proceedings of the National Academy of Sciences*, *114*(49), 12916–12921. https://doi.org/10.1073/pnas.1712966114

Bergelson, E., & Swingley, D. (2012). At 6 to 9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109*(9), 3253–3258. https://doi.org/10.1073/pnas.1113380109

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. *Cognition*, *127*(3), 391–397. https://doi.org/10.1016/j.cognition.2013.02.011

Bergelson, E., & Swingley, D. (2015). Early word comprehension in infants: Replication and extension. *Language Learning and Development*, *11*(4), 369–380. https://doi.org/10.1080/15475441.2014.979387

Bjorklund, D. F. (1995). *Children's thinking*. Brooks/Cole.

Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental Review*, *10*(1), 48–71. https://doi.org/10.1016/0273-2297(90)90004-N

Blanken, G. (1998). Lexicalisation in speech production: Evidence from form-related word substitutions in aphasia. *Cognitive Neuropsychology*, *15*(4), 321–360. https://doi.org/10.1080/026432998381122

Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access in language production. *Journal of Memory and Language*, *48*(3), 468–488. https://doi.org/10.1016/S0749-596X(02)00503-X

Brysbaert, M., & Ghyselinck, M. (2007). The effect of age of acquisition: Partly frequency related, partly frequency independent. *Visual Cognition*, *13*(7–8), 992–1011. https://doi.org/10.1080/13506280544000165

Brysbaert, M., & New, B. (2009). Moving beyond kučera and francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Budd, M.-J., Hanley, R., & Griffiths, Y. (2011). Simulating children's retrieval errors in picture-naming: A test of foygel and dell's (2000) semantic/phonological model of speech production. *Journal of Memory and Language*, *64*(1), 74–87. https://doi.org/10.1016/j.jml.2010.08.005

Butterfield, G. B., & Butterfield, E. C. (1977). Lexical codability and age. *Journal of Verbal Learning and Verbal Behavior*, *16*(1), 113–118. https://doi.org/10.1016/S0022-5371(77)80013-3

Butterworth, B. (1981). Speech errors: Old data in search of new theories. *Linguistics*, *19*(7–8), 627–662. https://doi.org/10.1515/ling.1981.19.7-8.627

Butterworth, B. (1989). Lexical access in speech production. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 108–135). MIT Press.

Butterworth, B. (1992). Disorders of phonological encoding. *Cognition*, *42*(1–3), 261–286. https://doi.org/10.1016/0010-0277(92)90045-J

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, *14*(1), 177–208. https://doi.org/10.1080/026432997381664

Carroll, J. B., & White, M. N. (1973). Word frequency and age-of-acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology*, *25*(1), 85–95. https://doi.org/10.1080/14640747308400325

Carter, C. S., Mintun, M., & Cohen, J. D. (1995). Interference and facilitation effects during selective attention: An H2150 PET study of stroop task performance. *NeuroImage*, *2*(4), 264–272. https://doi.org/10.1006/nimg.1995.1034

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 73–96). Lawrence Erlbaum.

Chinese Knowledge Information Processing Group. (1997). *Academia sinica balanced corpus* (WWW Version 3.0) Institute of Information Science, Academia Sinica.

Collins, A. M., & Loftus, E. F. (1975). Spreading activation theory of semantic processing. *Psychological Review*, *82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407

Comalli Jr., P. E., Wapner, S., & Werner, H. (1962). Interference effects of stroop colour-word test in childhood, adulthood, and aging. *The Journal of Genetic Psychology*, *100*(1), 47–53. https://doi.org/10.1080/00221325.1962.10533572

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(5), 1283–1296. https://doi.org/10.1037/0278-7393.26.5.1283

Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic Bulletin & Review*, *24*(4), 1158–1170. https://doi.org/10.3758/s13423-016-1191-6

Cowan, N., Fristoe, N. M., Elliott, E. M., Brunner, R. P., & Saults, J. S. (2006). Scope of attention, control of attention, and intelligence in children and adults. *Memory & Cognition*, *34*(8), 1754–1768. https://doi.org/10.3758/BF03195936

Cutting, J. C., & Ferreira, V. S. (1999). Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 318–344. https://doi.org/10.1037/0278-7393.25.2.318

Cycowicz, Y. M., Friedman, D., & Rothstein, M. (1997). Picture naming by young children: Norms for name agreement, familiarity, and visual complexity. *Journal of Experimental Child Psychology*, *65*(2), 171–237. https://doi.org/10.1006/jecp.1996.2356

Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture-word interference tasks. *Psychonomic Bulletin & Review*, *10*(1), 111–117. https://doi.org/10.3758/BF03196474

Damian, M. F., & Martin, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(2), 345–361. https://doi.org/10.1037/0278-7393.25.2.345

D'Amico, S., Devescovi, A., & Bates, E. (2001). Picture naming and lexical access in Italian children and adults. *Journal of Cognition and Development*, *2*(1), 71–105. https://doi.org/10.1207/S15327647JCD0201_4

Dawson, M. R. (1988). Fitting the ex-Gaussian equation to reaction time distributions. *Behavior Research Methods, Instruments, & Computers*, *20*(1), 54–57. https://doi.org/10.3758/BF03202603

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283

Dell, G. S., & O'Seaghdha, P. G. (1991). Mediated and convergent lexical priming in language production: A comment on Levelt et al. (1991). *Psychological Review*, *98*(4), 604–614. https://doi.org/10.1037/0033-295X.98.4.604

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, *42*(1–3), 287–314. https://doi.org/10.1016/0010-0277(92)90046-K

Dell, G. S., & Reich, P. A. (1981). Stages in sentence production: An analysis of speech error data. *Journal of Verbal Learning and Verbal Behavior*, *20*(6), 611–629. https://doi.org/10.1016/S0022-5371(81)90202-4

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

De Mauro, T., Mancini, F., Vedovelli, M., & Voghera, M. (1993). *Lessico di frequenza dell'italiano parlato* [Frequency lexicon of spoken Italian]. ETASLIBRI.

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. *Psychological Bulletin*, *89*(1), 63–100. https://doi.org/10.1037/0033-2909.89.1.63

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review*, *12*(1), 45–75. https://doi.org/10.1016/0273-2297(92)90003-K

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). Multipic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816. https://doi.org/10.1080/17470218.2017.1310261

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98–116. https://doi.org/10.1037/0012-1649.42.1.98

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228–231. https://doi.org/10.1111/1467-9280.00044

Ferreira, V., & Pashler, H. (2002). Central bottleneck influences on the processing stages of word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1187–1199. https://doi.org/10.1037/0278-7393.28.6.1187

Finocchiaro, C., & Caramazza, A. (2006). The production of pronominal clitics: Implications for theories of lexical access. *Language and Cognitive Processes*, *21*(1-3), 141–180. https://doi.org/10.1080/01690960400001887

Fox, J., & Weisberg, S. (2018). Visualizing fit and lack of fit in complex regression models with predictor effect plots and partial residuals. *Journal of Statistical Software*, *87*(9), 1–27.

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (3rd ed.). https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Foygel, D., & Dell, G. S. (2000). Models of impaired lexical access in speech production. *Journal of Memory and Language*, *43*(2), 182–216. https://doi.org/10.1006/jmla.2000.2716

Friedmann, N., Biran, M., & Dotan, D. (2013). Lexical retrieval and its breakdown in aphasia and developmental language impairment. In C. Boeckx & K. K. Grohmann (Eds.), *The Cambridge handbook of biolinguistics* (pp. 350–374). Cambridge University Press.

Füredi, M., & Kelemen, J. (1989). *A mai magyar nyelv szépprózai gyakorisgi szótra 1965–1977* [Prose frequency dictionary of contemporary Hungarian language 1965–1977]. Akadémiai Kiadó.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Language production. Vol. 1. Speech and talk* (pp. 177–220). Academic Press.

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*(2), 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, *35*(3), 515. https://doi.org/10.1017/S0305000907008641

Goulden, R., Nation, P., & Read, J. (1990). How large can receptive vocabulary be? *Applied Linguistics*, *11*(4), 341–363. https://doi.org/10.1093/applin/11.4.341

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, *82*(1), B1–B14. https://doi.org/10.1016/S0010-0277(01)00138-X

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word recognition. *Journal of Memory and Language*, *38*(3), 313–338. https://doi.org/10.1006/jmla.1997.2547

Guttentag, R. E., & Haith, M. M. (1978). Automatic processing as a function of age and Reading ability. *Child Development*, *49*(3), 707–716. https://doi.org/10.2307/1128239

Hardy, M. A. (1993). *Regression with dummy variables*. Sage Publications, Inc.

Harley, T. A. (1984). A critique of top-down independent levels models of speech production: Evidence from non-plan-internal speech errors. *Cognitive Science*, *8*(3), 191–219. https://doi.org/10.1207/s15516709cog0803_1

Harley, T. A. (1993). Phonological activation of semantic competitors during lexical access in speech production. *Language and Cognitive Processes*, *8*(3), 291–309. https://doi.org/10.1080/01690969308406957

Hohle, R. H. (1965). Inferred components of reaction times as functions of foreperiod duration. *Journal of Experimental Psychology*, *69*(4), 382–386. https://doi.org/10.1037/h0021740

Humphreys, K., Boyd, C., & Watter, S. (2010). Phonological facilitation from pictures in a word association task: Evidence for routine cascaded processing in spoken word production. *Quarterly Journal of Experimental Psychology*, *63*(12), 2289–2296. https://doi.org/10.1080/17470218.2010.509802

Jaeger, J. J. (2005). *Kid's slips: What young children's slips of the tongue reveal about language*. Lawrence Erlbaum Associates.

Jerger, S., Martin, R., & Pirozzolo, F. (1988). A developmental study of the auditory stroop effect. *Brain and Language*, *35*(1), 86–104. https://doi.org/10.1016/0093-934X(88)90102-2

Jerger, S., Martin, R. C., & Damian, M. F. (2002). Semantic and phonological influences on picture naming by children and teenagers. *Journal of Memory and Language*, *47*(2), 229–249. https://doi.org/10.1016/S0749-596X(02)00002-5

Jerger, S., Pearson, D., & Spence, M. (1999). Developmental course of auditory processing interactions: Garner interference and simon interference. *Journal of Experimental Child Psychology*, *74*(1), 44–67. https://doi.org/10.1006/jecp.1999.2504

Jescheniak, J. D., Hahne, A., Hoffman, S., & Wagner, C. (2006). Phonological activation of category coordinates during speech planning is observable in children but not in adults: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(2), 373–386. https://doi.org/10.1037/0278-7393.32.3.373

Jescheniak, J. D., Hahne, A., & Schriefers, H. (2003). Information flow in the mental lexicon during speech planning: Evidence from event-related brain potentials. *Cognitive Brain Research*, *15*(3), 261–276. https://doi.org/10.1016/S0926-6410(02)00198-2

Jescheniak, J. D., Hantsch, A., & Schriefers, H. (2005). Context effects on lexical choice and lexical activation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(5), 905–920. https://doi.org/10.1037/0278-7393.31.5.905

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824–843. https://doi.org/10.1037/0278-7393.20.4.824

Jescheniak, J. D., & Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: Further evidence from the coactivation of near-synonyms. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(5), 1256–1274. https://doi.org/10.1037/0278-7393.24.5.1256

Jescheniak, J. D., & Schriefers, H. (2001). Priming effects from phonologically related distractors in picture–word interference. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *54A*(2), 371–382.

Johnson, C. J. (1992). Cognitive components of naming in children: Effects of referential uncertainty and stimulus realism. *Journal of Experimental Child Psychology*, *53*(1), 24–44. https://doi.org/10.1016/S0022-0965(05)80003-7

Johnson, C. J., & Clark, J. M. (1988). Children's picture naming difficulty and errors: Effects of age of acquisition, uncertainty, and name generality. *Applied Psycholinguistics*, *9*(4), 351–365. https://doi.org/10.1017/S0142716400008055

Johnson, C. J., Paivio, A., & Clark, J. M. (1996). Cognitive components of picture naming. *Psychological Bulletin*, *120*(1), 113–139. https://doi.org/10.1037/0033-2909.120.1.113

Juhasz, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, *131*(5), 684–712. https://doi.org/10.1037/0033-2909.131.5.684

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming: Indirect election of words. *Cognition*, *14*(2), 185–209. https://doi.org/10.1016/0010-0277(83)90029-X

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology*, *25*(4), 463–492. https://doi.org/10.1080/02643290701674851

Kleiman, E. (2021). *EMAtools: Data management tools for real-time monitoring/ecological momentary assessment data*. R package version 0.1.4. https://CRAN.R-project.org/package = EMAtools

Kuipers, J.-R., & La Heij, W. (2009). The limitations of cascading in the speech production system. *Language and Cognitive Processes*, *24*(1), 120–135. https://doi.org/10.1080/01690960802234177

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). Lmertest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lachman, R. (1973). Uncertainty effects on time to access the internal lexicon. *Journal of Experimental Psychology*, *99*(2), 199–208. https://doi.org/10.1037/h0034633

Lachman, R., & Lachman, J. L. (1980). Picture naming: Retrieval and activation of long-term memory. In L. W. Poon, J. L. Fozard, L. S. Cermak, D. Arenberg, & L. W. Thompson (Eds.), *New directions in memory and aging* (pp. 313–343). Erlbaum.

Lachman, R., Shaffer, J. P., & Hennrikus, D. (1974). Language and cognition: Effects of stimulus codability, name-word frequency, and age of acquisition on lexical reaction time. *Journal of Verbal Learning and Verbal Behavior*, *13*(6), 613–625. https://doi.org/10.1016/S0022-5371(74)80049-6

La Heij, W., Dirkx, J., & Kramer, P. (1990). Categorical interference and associative priming in picture naming. *British Journal of Psychology*, *81*(4), 511–525. https://doi.org/10.1111/j.2044-8295.1990.tb02376.x

Lee, E.-K., Brown-Schmidt, S., & Watson, D. G. (2013). Ways of looking ahead: Hierarchical planning in language production. *Cognition*, *129*(3), 544–562. https://doi.org/10.1016/j.cognition.2013.08.007

Lenth, R. (2019). *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.3.4. https://CRAN.R-project.org/package = emmeans

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. MIT Press.

Levelt, W. J. M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of*

Sciences, *98*(23), 13464–13471. https://doi.org/10.1073/pnas.231459498

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

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological Review*, *98*(1), 122–142. https://doi.org/10.1037/0033-295X.98.1.122

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization*. Oxford University Press.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Lawrence Erlbaum Associates.

Mädebach, A., Jescheniak, J. D., Oppermann, F., & Schriefers, H. (2011). Ease of processing constrains the activation flow in the conceptual-lexical system during speech panning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 649–660. https://doi.org/10.1037/a0022330

Mahon, B., Costa, A., Peterson, R., Vargas, K. A., & Caramazza, A. (2007). Lexical selection is not by competition: A reinterpretation of semantic interference and facilitation effects in the picture–word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(3), 503–535. https://doi.org/10.1037/0278-7393.33.3.503

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake – But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(4), 843–847. https://doi.org/10.1037/a0029284

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS One*, *7*(8), e43230. https://doi.org/10.1371/journal.pone.0043230

Martin, N., Gagnon, D. A., Schwartz, M. F., Dell, G. S., & Saffran, E. M. (1996). Phonological facilitation of semantic errors in normal and aphasic speakers. *Language and Cognitive Processes*, *11*(3), 257–282. https://doi.org/10.1080/016909696387187

Martin, N., Weisberg, R. W., & Saffran, E. M. (1989). Variables influencing the occurrence of naming errors: Implications for models of lexical retrieval. *Journal of Memory and Language*, *28*(4), 462–485. https://doi.org/10.1016/0749-596X(89)90022-3

Massidda, D. (2013). *retimes: Reaction time analysis*. R package version 0.1-2. https://CRAN.R-project.org/package = retimes

McAuliffe, M., Socolof, M., Mihuc, S., Wagner, M., & Sonderegger, M. (2017). Montreal Forced Aligner: trainable text-speech alignment using Kaldi. *Proceedings of the 18th Conference of the International Speech Communication Association*, 498–502.

Melinger, A., & Abdel Rahman, R. (2013). Lexical selection is competitive: Evidence from indirectly activated semantic associates during picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(2), 348–364. https://doi.org/10.1037/a0028941

Meyer, A. S., & Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Memory & Cognition*, *35*(3), 494–503. https://doi.org/10.3758/BF03193289

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(6), 1146–1160. https://doi.org/10.1037/0278-7393.17.6.1146

Mills, D. L., Prat, C., Zangl, R., Stager, C. L., Neville, H. J., & Werker, J. F. (2004). Language experience and the organization of brain activity to phonetically similar words: ERP evidence from 14- and 20-month olds. *Journal of Cognitive Neuroscience*, *16*(8), 1–13. https://doi.org/10.1162/0898929042304697

Momma, S. (2021). Filling the gap in gap-filling: Long-distance dependency formation in sentence production. *Cognitive Psychology*, *129*, 101411. https://doi.org/10.1016/j.cogpsych.2021.101411

Morsella, E., & Miozzo, M. (2002). Evidence for a cascade model of lexical access in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(3), 555–563. https://doi.org/10.1037/0278-7393.28.3.555

Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, *20*(2), 233–253. https://doi.org/10.2307/747758

Navarette, E., Peressotti, F., Lerose, L., & Miozzo, M. (2017). Activation cascading in sign production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(2), 302–318. https://doi.org/10.1037/xlm0000312

Navarrete, E., & Costa, A. (2009). The distractor picture paradox in speech production: Evidence from the word translation task. *Journal of Psycholinguistic Research*, *38*(6), 527–547. https://doi.org/10.1007/s10936-009-9119-1

Oldfield, R. C., & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology*, *17*(4), 273–281. https://doi.org/10.1080/17470216508416445

Paivio, A., Clark, J. M., Digdon, N., & Bons, T. (1989). Referential processing: Reciprocity and correlates of naming and imaging. *Memory & Cognition*, *17*(2), 163–174. https://doi.org/10.3758/BF03197066

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(3), 539–557. https://doi.org/10.1037/0278-7393.24.3.539

Poarch, G. J., & van Hell, J. G. (2012). Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, *111*(3), 419–438. https://doi.org/10.1016/j.jecp.2011.09.008

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*(3), 460–499. https://doi.org/10.1037/0033-295X.107.3.460

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*(3), 446–461. https://doi.org/10.1037/0033-2909.86.3.446

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114*(3), 510–532. https://doi.org/10.1037/0033-2909.114.3.510

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Ridderinkhof, K. R. (2002). Activation and suppression in conflict tasks: Empirical clarification through distributional analyses. In W. Prinz & B. Hommel (Eds.), *Attention and performance XIX: Common mechanisms in perception and action* (pp. 494–519). Oxford University Press.

Riggs, K. J., Mctaggart, J., Simpson, A., & Freeman, R. P. (2006). Changes in the capacity of visual working memory in 5- to

10-year-olds. *Journal of Experimental Child Psychology*, *95*(1), 18–26. https://doi.org/10.1016/j.jecp.2006.03.009

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, *42*(1-3), 107–142. https://doi.org/10.1016/0010-0277(92)90041-F

Roelofs, A. (2008). Tracing attention and the activation flow of spoken word planning using eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 353–368. https://doi.org/10.1037/0278-7393.34.2.353

Roelofs, A., Meyer, A. S., & Levelt, W. J. M. (1996). Interaction between semantic and orthographic factors in conceptually driven naming: Comment of starreveld and LaHeij (1995). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(1), 246–251. https://doi.org/10.1037/0278-7393.22.1.246

Rossion, B., & Pourtois, G. (2004). Revisiting snodgrass and vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*(2), 217–236. https://doi.org/10.1068/p5117

Sadat, J., Martin, C. D., Costa, A., & Alario, F.-X. (2014). Reconciling phonological neighborhood effects in speech production through single trial analysis. *Cognitive Psychology*, *68*, 33–58. https://doi.org/10.1016/j.cogpsych.2013.10.001

Schneider, W., & Bjorklund, D. F. (1998). Memory. In W. Damon, D. Kuhn, & R. S. Siegler (Eds.), *Handbook of child psychology/ Vol. 2, cognitive, language, and perceptual development* (5th ed., pp. 467–521). Wiley.

Schriefers, H., Meyer, A. S., & Levelt, W. J. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. *Journal of Memory and Language*, *29*(1), 86–102. https://doi.org/10.1016/0749-596X(90)90011-N

Schriefers, H., & Vigliocco, G. (2015). Speech production, psychology of. In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences* (2nd ed., pp. 255–258). Elsevier.

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, *54*(2), 228–264. https://doi.org/10.1016/j.jml.2005.10.001

Shipley, K. G., & McAfee, J. G. (2015). *Assessment in speech-language pathology: A resource manual* (5th Ed.). Cengage Learning.

Sieger-Gardner, L., & Schwartz, R. G. (2008). Lexical access in children with and without specific language impairment: A cross–model picture–word interference study. *International Journal of Language & Communication Disorders*, *43*(5), 528–551. https://doi.org/10.1080/13682820701768581

Simmering, V. R. (2012). The development of visual working memory capacity during early childhood. *Journal of Experimental Child Psychology*, *111*(4), 695–707. https://doi.org/10.1016/j.jecp.2011.10.007

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*(2), 174–215. https://doi.org/10.1037/0278-7393.6.2.174

Spieler, D. H., & Griffin, Z. M. (2006). The influence of age on the time course of word preparation in multiword utterances. *Language and Cognitive Processes*, *21*(1-3), 291–321. https://doi.org/10.1080/01690960400002133

Stafford, T., & Gurney, K. N. (2011). Additive factors do not imply discrete processing stages: A worked example using models of the stroop task. *Frontiers in Psychology*, *2*, Article 287. https://doi.org/10.3389/fpsyg.2011.00287

Starreveld, P. A. (2000). On the interpretation of onsets of auditory context effects in word production. *Journal of Memory and Language*, *42*(4), 497–525. https://doi.org/10.1006/jmla.1999.2693

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(3), 686–698. https://doi.org/10.1037/0278-7393.21.3.686

Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. *Cognition*, *114*(3), 447–454. https://doi.org/10.1016/j.cognition.2009.11.003

Stemberger, J. P. (1983). Inflectional malapropisms: Form based errors in English morphology. *Linguistics*, *21*, 573–602. https://doi.org/10.1515/ling.1983.21.4.573

Sternberg, S. (1969). The discovery of processing stages: Extensions of donders' method. *Acta Psychologica*, *30*, 276–315. https://doi.org/10.1016/0001-6918(69)90055-9

Sternberg, S. (1984). Stage models of mental processing and the additive-factor method. *Behavioral and Brain Sciences*, *7*(1), 55–94. https://doi.org/10.1017/S0140525X00026285

Sternberg, S. (1998). Discovering mental processing stages: The method of additive factors. In D. Scarborough & S. Sternberg (Eds.), *An invitation to cognitive science: Methods, models, and conceptual issues* (Vol. 4, pp. 703–863). MIT Press.

Sternberg, S. (2001). Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychologica*, *106*(1–2), 147–246. https://doi.org/10.1016/S0001-6918(00)00045-7

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, *20*(4), 912–928. https://doi.org/10.1093/cercor/bhp153

Sylvia, L. (2017). *Lexical access in children with specific language impairment and deficits in auditory processing*. [Doctoral dissertation, City University of New York]. CUNY Academic Works.

Székely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., Arévalo, A., & Bates, E. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research Methods, Instruments, & Computers*, *35*(4), 621–633. https://doi.org/10.3758/BF03195542

Székely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., Arévalo, A., Vargha, A., & Bates, E. (2005). Timed action and object naming. *Cortex*, *41*(1), 7–25. https://doi.org/10.1016/S0010-9452(08)70174-6

Thomas, R. D. (2006). Processing time predictions of current models of perception in the classic additive factors paradigm. *Journal of Mathematical Psychology*, *50*(5), 441–455. https://doi.org/10.1016/j.jmp.2006.05.006

Tincoff, R., & Juscyzk, P. W. (1999). Some beginnings of word comprehension in 6-month-olds. *Psychological Science*, *10*(2), 172–175. https://doi.org/10.1111/1467-9280.00127

Tincoff, R., & Juscyzk, P. W. (2011). Six-month-olds comprehend words that refer to parts of the body. *Infancy*, *17*(4), 432–444. https://doi.org/10.1111/j.1532-7078.2011.00084.x

Vigliocco, G., Vinson, D. P., Lewis, W., & Garrett, M. (2004). Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology*, *48*(4), 422–488. https://doi.org/10.1016/j.cogpsych.2003.09.001

Vincent, S. B. (1912). The function of the viborissae in the behavior of the white rat. *Behavioral Monographs*, *1*(5), 81.

Vurpillot, E., & Ball, W. A. (1979). The concept of identity and children's selective attention. In G. A. Hale & M. Lewis (Eds.), *Attention and cognitive development* (pp. 23–42). Plenum.

Werker, J. F., Fennel, C. T., Corcoran, K. M., & Stager, C. L. (2002). Infants' ability to learn phonetically similar words: Effects of age and vocabulary size. *Infancy*, *3*(1), 1–30. https://doi.org/10.1207/S15327078IN0301_1

Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development*, *6*(2), 179–208. https://doi.org/10.1207/s15327647jcd0602_2

Zareva, A., Schwanenflugel, P., & Nikolova, Y. (2005). Relationship between lexical competence and language proficiency: Variable sensitivity. *Studies in Second Language Acquisition*, *27*(4), 567–596. https://doi.org/10.1017/S0272263105050254

Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word Reading and other tasks. *Journal of Memory and Language*, *47*(1), 1–29. https://doi.org/10.1006/jmla.2001.2834

Zhang, H., Carlson, M. T., & Diaz, M. T. (2020). Investigating effects of phonological neighbors on word retrieval and phonetic variation in word naming and picture naming paradigms. *Language, Cognition, and Neuroscience*, *35*(8), 980–991. https://doi.org/10.1080/23273798.2019.1686529

Zhang, Q., Zhu, X., & Damian, M. (2018). Phonological activation of category coordinates in spoken word production: Evidence for cascaded processing in English but not in mandarin. *Applied Psycholinguistics*, *39*(5), 835–860. https://doi.org/10.1017/S0142716418000024