Journal of Memory and Language

On idle idols and ugly icons: Investigating lexical selection in typing through homophones

Merel Muylle[1], Robert J. Hartsuiker[1], and Nazbanou Nozari[2, 3]

[1]Department of Experimental Psychology, Ghent University, Henri Dunantlaan 2, 9000 Ghent, Belgium

[2]Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10th St., Bloomington, IN 47405-7007, United States

[3] Cognitive Science Program, 1101 E. 10th St., Bloomington, IN 47405-7007, United States

E-mail

merel.muylle@ugent.be

bnozari@iu.edu

robert.hartsuiker@ugent.be

Funding: This work was supported by a grant from Ghent University Special Research Fund to MM, and NSF grant BCS-2317121 to NN.

Corresponding author

Merel Muylle

Ghent University, Department of Experimental Psychology Henri Dunantlaan 2

9000 Ghent

E-mail: merel_muylle@hotmail.com

Abstract

Homophone errors (e.g., *there/their*) are not uncommon in typing, but it is debated whether they simply reveal poor spelling knowledge or signal competition in the production system. We tested the idea that competition underlies the greater difficulty associated with producing homophone, compared to non-homophone, targets. Using computational simulations, we showed that competition alone is sufficient to produce interference during homophone production, and that such interference is exacerbated by increasing lexical competition. These predictions were confirmed in two experiments, a typing-to-dictation task (Experiment 1) and a question-answering task (Experiment 2). We further showed the homophone effect was insensitive to the syntactic category: we found a robust homophone interference effect of a comparable magnitude for same-category (e.g., flower-flour) and different-category (e.g., idol-idle) homophones. Collectively, these results show that lexical selection in typing is similar to speaking in terms of the processes arising from representational overlap, but distinct from it in terms of the influence of syntax.

Keywords: typing; homophones; sentence production

Introduction

Most of us have experienced replacing words with a (heterographic) homophone (e.g., *there*, *they're*, and *their*) when writing or typing. According to some researchers, homophone substitution errors are not very common, since the slower pace of typing/writing (compared to speech) allows for correcting these errors before or immediately after they emerge (Hotopf, 1980, 1983; Largy, 1996; Morton, 1980). Nevertheless, analyses of writing samples have shown that homophone substitution errors are the most common misspellings that end up in final texts (the top two most observed errors *to/too* and *there/their* were homophone substitutions in untimed essays by college students; Connors & Lunsford, 1992; about 25% of misspellings in 10-minute short essays by university students; Lastres López & Manalastas, 2018). Yet, the origin of these errors remains obscure.

Undoubtedly, homophone substitution errors sometimes originate from poor spelling knowledge (Bonin et al., 2001). However, it is unclear whether competition in the linguistic system, per se and in the absence of poor spelling knowledge, can cause homophone errors. Although the idea has been discussed in the literature (e.g., Morton, 1980; White et al., 2008), the evidence in its support has been circumstantial, as we will discuss below. This paper aims to fill this gap. Specifically, we hypothesize that the co-activation of the linguistic representations induced by phonological overlap in homophones (e.g., /sʌn/ for *sun* and *son*) induces lexical competition, which makes homophone target production both slower and more error-prone than non-homophone target production. We show that this "homophone interference effect" persists even after ruling out spelling knowledge errors and is directly related to the degree of lexical competition in the system. We then test syntactic constraints on homophone production to further elucidate the nature of homophone representations.

Current evidence for the source of the homophone interference effect

One of the most detailed studies on homophones was conducted by Bonin, Peereman, and Fayol (2001). In a timed handwritten picture naming task, the authors compared the onset reaction times (RTs) and error rates of French subordinate homophones (e.g., *ver* [worm]), which have a more frequent alternative (e.g., *verre* [glass]), to control words matched on frequency and phoneme-to-grapheme consistency (i.e., how many sound-to-letter mappings were possible for the involved phonemes). Homophones were written with more errors, but not more slowly. However, the difference in accuracy disappeared with spelling knowledge as a covariate in the analysis. The absence of differences in timing, together with higher error rates on words that failed a spelling knowledge test, led the authors to conclude that poor spelling knowledge was the primary source of homophone errors. However, in that study, spelling knowledge was tested on an independent group of participants, and not the original participants. This is problematic given the evidence for individual differences in spelling knowledge (Kamhi & Hinton, 2000), leaving open the possibility that Bonin et al. (2001) may have missed processing difficulties that cannot be attributed to a spelling knowledge deficit at the individual level.

In contrast, several researchers have proposed that the homophone interference effect is not merely a consequence of poor spelling knowledge, but a natural product of the inner dynamics of the language production system (e.g., Morton, 1980; White et al., 2008). Partial motivation for this claim comes from studies in perception, which found more errors and longer RTs in the categorization and reading aloud of homophones (e.g., Biedermann et al., 2009; Coltheart et al., 1994; van Orden, 1987). However, reading entails visual decoding, which could be a unique source of confusion, absent in many production tasks. At the same time, there is more direct evidence from production tasks that also motivates competition as a direct source of the homophone interference effect. When representations overlap in meaning or form, they activate one another, leading to increased competition for selection (e.g., cat/dog: Damian et al., 2001; Kroll & Stewart, 1994; Nozari et al., 2016; Schnur, 2014; or cat/mat: Breining et al., 2016; McCloskey et al., 2006; Pinet & Nozari, 2018; Romani et al., 2002; Wheeldon, 2003). While theories of language production differ in their mechanistic explanation of interference effects (e.g., Oppenheim et al., 2010; Oppenheim & Nozari, 2024; Roelofs, 1992; Roelofs, 2018), most of them agree that competition makes production slower and more error prone (Cutting & Ferreira, 1999; Dell, 1986; Levelt et al., 1999; Plaut & Shallice, 1993; Romani et al., 2011; Stemberger, 1982; cf., Dhooge & Hartsuiker, 2010; Mahon et al., 2007; Navarrete et al., 2014).

In the case of homophones, two words with different meanings and spellings overlap in phonology (e.g., /sʌn/ for *sun* and *son*). Although form overlap can elicit facilitation in speech (see Goldrick et al., 2010, for a review) and writing (see Bonin et al., 2015, for a review), especially when the overlap is predictable, similarity in form also elicits robust interference in spoken, handwritten, and typed production (Breining et al., 2016, 2019; Nozari et al., 2016; Pinet & Nozari, 2018, 2023; Rogers & Storkel, 1998; Sadat et al., 2014; Sevald & Dell, 1994; Sullivan, 1999; Wheeldon, 2003). For instance, in a blocked cyclic naming paradigm, Breining and colleagues (2016) showed that participants were slower to name and write words, when they had form overlap with the other words within the block (e.g., *pot, peg, leg, log, pig, pill*) than when they did not. In a later study, Harrison, Hepner, and Nozari (2020) showed identical patterns of segmental interference in handwriting and typing modalities. Similarly, pairs of words that share a vowel vs. those that do not (e.g., *fog top* vs. *fig top*) elicit more errors on the consonants both in spoken (Dell, 1984) and typed (Pinet & Nozari, 2018) production. Finally, cognates, i.e., similar-sounding words across languages (e.g., *apple/Apfel*), can cause interference in spoken (Martin & Nozari, 2021) and typed production (Muscalu & Smiley, 2019; Muylle et al., 2022, see Bailey et al., 2024, for a review of facilitation and interference effects in cognates).

Thus, by the competition logic, the phonological overlap in homophones should naturally lead to interference, even in the absence of spelling knowledge deficits. A few studies have approached this question via different manipulations. We discuss these findings to highlight the important insights they offer, but also to show that they have not yet provided convincing evidence for competition as a core mechanism underlying the homophone interference effect.

White et al. (2008) investigated homophone errors in a handwritten sentence dictation task and found that producing the non-dominant (i.e., lower frequency) homophone (e.g., *beech*) resulted in more errors than producing the dominant (i.e., higher frequency) homophone (e.g., *beach*). Since the authors excluded the items that were incorrectly identified in a homophone recognition task, this frequency effect cannot be attributed to a lack of spelling knowledge for less frequent words. In addition, the presentation of a word that primed the non-target spelling (e.g., *teacher* for the target *beech*) earlier in the sentence resulted in an increase of homophone errors, and equally so for dominant and non-dominant targets. Follow-up studies replicated both frequency and priming effects on homophone writing in younger and older populations (White et al., 2013) and in typing as an output modality (White et al., 2010). Finally, White et al. (2012) found semantic priming effects on homophone typing, with fewer homophone errors when the target meaning was primed (e.g., *wind* for the target *blew*) and more homophone errors when the non-target meaning was primed (e.g., *sky* for the target *blew*) compared to an unrelated prime. This finding suggests that lexical competition can be induced by priming the meaning of the homophone competitor.

One caveat in the studies of White et al. (2008, 2013) is that they did not control for phonological-orthographic (PO) consistency across dominant and subordinate homophones (but see White et al., 2012, who also studied homophone regularity effects). Some PO mappings are more frequent than others (e.g., the rime -/iʧ/ is more frequently spelled as -*each* than as -*eech*, see Chee et al., 2020). If the more frequent PO mapping happens to be more frequent in the dominant homophone, the locus of the dominance effect may still be sublexical. This possibility matches with the absence of an interaction between the frequency effect and priming, which also points to a sublexical locus of the dominance effect. Second, there are no timing data. Without these, the findings would still not provide convincing evidence for lexical competition as the underlying cause of the homophone interference effect, even if the PO distribution had been balanced across frequencies. Probabilistically speaking, the dominant representation could be automatically activated in the majority of trials, with the subordinate representation having near-zero activation, especially since spelling-to-dictation primarily activates phonology and to a lesser extent meaning. Such a situation would generate a dominance effect identical to that shown in the above studies, without needing to propose any competition between two representations. The scenario described above predicts full homophone substitutions, such as the errors observed by White et al. (2008), but it would not predict a slow-down in homophone production. On the other hand, if two representations are both active and competing for selection, most selection theories propose a slowdown (Roelofs, 2018; Nozari, 2025). Therefore, timing data are essential in posing competition as the mechanism underlying the homophone interference effect. Importantly, to interpret timing data, it is necessary to add matched control words to the design.

To summarize, the extant literature on the co-activation of lexical representations in language production suggests that competition should be a core mechanism underlying homophone production. However, demonstrating so requires showing that homophones (which have a built-in lexical competitor) are produced more slowly than well-matched controls (which do not have a built-in lexical competitor). Of the studies reviewed above, only Bonin et al. (2001) compared production time between homophones and matched controls, and reported a null effect on this measure, which may be due to any number of reasons, such as low statistical power. The other studies did not match their homophones on PO consistency and did not provide timing data that can test a competition account of homophone interference. The current study was designed to address these problems.

What is the nature of homophone representations?

When speaking of competition, it is important to consider what is the nature of the representations involved in competition. There are two general views on how homophones are lexically represented in the production system. These views differ on the number of *lemmas* and *lexemes* they attribute to homophones. Lemmas are lexical representations that are shared between all modalities of language production (e.g., Bock & Levelt, 1994). Lexemes are modality-specific lexical representations that differ between spoken and written/typed modalities (e.g., Bock & Levelt, 1994; Caramazza, 1997). Some researchers argue that homophone pairs have two separate lemmas but one lexeme (e.g., Biedermann et al., 2002; Jescheniak & Levelt, 1994). We call this the *double-lemma* view. Others claim that each homophone has its own lexeme, regardless of whether they accept the existence of a modality-independent lemma (e.g., Bonin & Fayol, 2002) or not (e.g., Caramazza et al., 2001; Caramazza & Miozzo, 1998; Jacobs et al., 2004; Miozzo & Caramazza, 2005). We call this the *double-lexeme* view.

The influence of syntax on homophone selection can be directly leveraged to test these opposing proposals. If there is only one lexeme, syntactic information must necessarily be stored at the lemma level (see also Bock & Levelt, 1994; Dell, 1986; Levelt et al., 1999), as homophones can belong to different syntactic categories (e.g., idol [noun] - idle [adjective]). Since, by definition, lemmas are shared with spoken production, then syntactic influences on homophone typing should mirror syntactic influences observed in spoken production. If, on the other hand, homophones can have their own independent lexemes, then syntax can be placed at the lexeme level, which is, by definition, separate from lexemes in spoken production. This view, thus, allows syntactic effects on homophone typing to be quite different from those observed in spoken production.

In spoken production, syntax strongly constrains lexical selection (e.g., Dell, 1986; Ferreira & Slevc, 2007; Levelt et al., 1999; Stemberger, 1982). Lexical errors in spoken production tend to respect syntactic categories (e.g., a noun tends to be replaced by a noun, and not a verb; Garrett, 1975, 1976). Also nouns with many noun neighbors are named more slowly than those with a comparable number of neighbors from a different grammatical class, and this difference is exaggerated in a sentence, compared to isolated noun, production (Heller & Goldrick, 2014). Correspondingly, some models of spoken production only allow the selection of lexical representations from the appropriate syntactic category (e.g., Dell, 1986). Similarly, trained models learn to use syntax to limit lexical selection (e.g., Gordon & Dell, 2003). However, evidence from orthographic domains, such as handwriting, tells a different story. For example, lexical errors violate syntactic categories much more often in handwriting than in speaking (Hotopf, 1980; Romani et al., 2002). More importantly, studies on aphasia have shown that grammatical deficits can be modality-specific (Caramazza & Hillis, 1991; Hillis & Caramazza, 1995; Rapp & Caramazza, 1997). These findings have given rise to accounts that place syntactic effects at the level of lexemes, rather than lemmas (e.g., Caramazza & Hillis, 1991; Starreveld & La Heij, 2004). Since lexemes in spoken and written/typed production are different for the same word, syntax can influence them differently, thus explaining the dissociations observed across modalities in their sensitivity to syntax.

If homophone errors in typing are strongly modulated by syntactic category constraints, as expected by a syntactic locus operating at the level of lemmas, the double-lemma view can be supported. Conversely, if homophone errors in typing show strong insensitivity to syntactic categories, similar to handwriting errors, the locus of syntax cannot be lemmas. This, in turn, means that at least two lexemes are required for homophone mates belonging to different syntactic categories, thus supporting the double-lexeme view.

White et al. (2010) tackled this question in a spelling-to-dictation task. Participants heard and typed sentences like "*Edward had to* ***explain/justify*** *to his mother why the broken* **pane** *on the window was an accident.*" and "*Amanda had met the* ***ideal/perfect*** *man, but she wasn't sure how to* **reel** *him in.*". Half of the target homophones belonged to the same syntactic category (e.g., pane - pain), while the other half belonged to different syntactic categories (e.g., reel - real). Moreover, the homophone target (e.g., *reel*) was preceded by a prime that either cued the spelling of the non-target homophone (e.g., *ideal*) or was unrelated to the homophones (e.g., *perfect*). Half of the primes belonged to the same syntactic category as the non-target homophone (e.g., ideal/real[1][adjective/adjective]), whereas the other half belonged to a different syntactic category (e.g., appeal/real [verb/adjective]). The authors found that homophone errors were overall more likely for the same-category compared to the different-category homophones. When the target and non-target homophones belonged to the same syntactic category, the prime type did not matter. However, when they belonged to different syntactic categories, only primes belonging to the same category as the competitor increased homophone substitution errors. Collectively, the authors took this pattern to imply that homophone production was constrained by syntax.

White et al. (2010) did not control for the frequency of same- and different-category homophones, which, given the past reports regarding the influence of lexical frequency on homophone production, could have affected the results. However, a more fundamental issue with their conclusion is the absence of non-homophone control words in similar sentences to compare to homophone targets. For instance, the current comparison is between "pane" in the same-category and "reel" in the different-category conditions, but the two words also differ in many other ways, including the length, CV structure, potential difficulty of their carrier sentences, and their positions within those sentences. These basic differences can easily cause differences in the processing of the target homophone that have nothing to do with the syntactic category of the non-target homophone. To test whether the syntactic category of the competing homophone truly constrains the processing of the target homophone, the design must contain control sentences with the matched non-homophone targets in the same position as the homophone targets (e.g., "*Edward had to justify to his mother why the broken* ***pane/rail*** *on the window was an accident.*" and "*Amanda had met the perfect man, but she wasn't sure how to* ***reel/draw*** *him in.*"). We need these controls to rule out that someone would type 'real' instead of 'reel' for reasons that are not due to the homophone state, but result

[1] Although *ideal* can also be used as a noun, it appeared as an adjective in the sentence in the experiment. Ideally, a clean separation of syntactic categories would utilize words that cannot possibly belong to the other category.

from postlexical processes (e.g., one may spell 'feel' as 'feal' based on PO conversion rules). Any difference between the target homophones "pane" and "reel" could then be interpreted as a syntactic effect after subtracting differences between the non-homophone controls "rail" and "draw".

Given these problems, it cannot be ruled out that the higher percentage of errors in same-category homophones was due to differences other than syntactic category overlap. More generally, White et al. (2010) may have overestimated syntactic influences on homophone selection when claiming that different-category homophone errors are infrequent and "most likely due to activation of incorrect orthography rather than incorrect lemma selection" (p. 166). For one thing, very common homophone errors often cross syntactic boundaries (e.g., to/too and there/their, Connors & Lundsford, 1992). Moreover, certain errors in homophone studies also point to syntactic violations. For example, Largy et al. (1996) examined homophonic noun-verb inflection errors on third-person singular verbs in French (e.g., *elle timbre* [she stamps]) vs. plural noun endings (e.g., *les timbres* [the stamps]) in a handwritten sentence dictation task. By using pronoun-pronoun-verb sentences (e.g., *elle les timbre* [she stamps them]), they took advantage of the local attraction of *les* (usually an article, but sometimes a pronoun), to prime the plural noun (*timbres*). This manipulation was successful showing that even within a sentence, homophone errors can cross syntactic boundaries. However, although not fitting with the third-person singular subject "elle", "timbres" can still be a verb (second-person singular). It is, therefore, necessary to verify this violation of syntactic categories using homophones that uniquely belong to one syntactic category or another. This was set as the last goal of the present study.

Present study

The overarching goal of the present study was to elucidate the nature of representations and processes involved in homophone production in typing. We break down this goal into three subgoals: (1) to assess competition as a core process underlying the homophone interference effect. Using the appropriate control conditions, we did not only study full homophone substitutions, but any mistakes that were made on the target word. The logic is that the chance of non-homophone segmental errors (e.g., finger slips or typos) should be the same in homophones and non-homophones. Thus, when comparing (pseudo)homophone errors in homophones and non-homophones that are matched on PO consistency, we can attribute the excess number of misspellings in homophones to the presence of the homophone competitor. Importantly, we also measured the production speed in various parts of the sentence to assess whether competition causes a slowdown. (2) To study lexical competition in a more direct way, we presented either one homophone or both homophones within the same sentence. Finally, (3) we tested the double-lemma vs. double-lexeme view by using a carefully controlled syntactic manipulation.

To achieve these subgoals, we used computational models to generate clearer predictions for empirical testing. Although several computational models of typing exist (e.g., Logan & Crump, 2011; Rumelhart & Norman, 1982), they typically focus on motor, rather than psycholinguistic, aspects of typing (see Pinet et al., 2016). Since our interest is in testing competition, we used the core architecture and mechanisms of the two-step interactive activation model of speech production (e.g., Dell, 1986), which has been successfully adapted to capture dynamics related to competition in speaking (e.g., Nozari et al., 2011; Burgess & Nozari, 2022). To the basic model, we added a syntactic and an orthographic layer to model a typing-to-dictation task. This model is well-suited for our purpose, as the spreading activation between semantics, lexical items, and segments (phonological and orthographic) naturally captures coactivation, a precursor of competition, and also allows for cascading and interactivity, which have been shown to exist in typing (Pinet & Nozari, 2018). Moreover, it allows us to pinpoint the level(s) at which homophone interference (i.e., lexical or orthographic) arises.

Simulation 1 demonstrated whether the homophone interference effect can be a natural consequence of competition in the system without assuming poor spelling knowledge. Simulation 2 tested the effect of increased lexical competition on the homophone interference effect by assuming the sentence contains both the target and the competing homophone. Simulation 3 varied the contribution of syntax continuously from very weak to very strong for homophones belonging to the same (Simulation 3a) vs. a different syntactic category (Simulation 3b), to assess the space of possibilities for the influence of syntax on homophone typing.

The predictions obtained from the three simulations were tested in two experiments. Experiment 1 was a typing-to-dictation task, similar to White et al. (2010). Experiment 2 was a question-answer typing paradigm, which was administered for two purposes: (a) replicating the findings with a different task, since most studies on typing have been focusing on spelling-to-dictation (Pinet et al., 2016), and (b) having a task that relies more heavily on the processing of meaning (i.e., participants have to understand the question and construct an answer). Apart from the differential involvement of meaning, the tasks also differ in cognitive load and preparation time, both of which may impact the results. The general design and manipulations were the same across the two experiments. Materials consisted of quadruplets with homophone (H) pairs and matched control (C) words under four conditions (see 1a-1d):

(1a) HH: The fashion idol turned out to be quite idle.

(1b) HC: The fashion idol turned out to be quite ugly.

(1c) CH: The fashion icon turned out to be quite idle.

(1d) CC: The fashion icon turned out to be quite ugly.

Half of the homophone pairs belonged to the same and the other half to a different syntactic category. For each target word, we measured accuracy and typing times. Concretely, we obtained three timing measures: onset reaction times (RT, i.e., first keypress of the target word), target durations (i.e., the time needed to complete the typing of the target, measured as the mean inter-keystroke interval or IKI), and pre-target duration (i.e., the time needed to type everything before the target, also in terms of mean IKI). Previous studies on homophones measured accuracy and/or onset reaction times, but none of them have looked into durations. Past studies have shown that the effect of phonological similarity sometimes manifests in RTs and sometimes in durations (Breining et al., 2016; Nozari et al., 2016; Pinet & Nozari, 2023), most likely reflecting differences in whether to prioritize a quick start or not (see Nozari & Hepner, 2019a). For this reason, we collected and analyzed all four measures.

We made the following predictions. First, we expect to replicate the findings of prior studies, i.e., the larger error rates on homophone targets than on controls. With regard to the three subgoals, (1) If homophone errors are a natural consequence of competition, processing should be slowed down for H, compared to C words, even when spelling knowledge is intact and the correct word is typed. (2) If increasing lexical competition increases the magnitude of the homophone interference effect, we would expect longer latencies and lower accuracy for H words in double homophone sentences (i.e., HH) than in single homophone sentences (i.e., HC and CH). Finally, (3) if syntax has a strong influence on competition that gives rise to the homophone interference effect, we would expect longer latencies and lower accuracy for H words belonging to the same vs. different syntactic categories. A strong syntax effect could be taken as support for the double-lemma view, without needing to pose double lexemes. Conversely, weak or non-existing effect of syntax in this analysis would support the double-lexeme view.

Model simulations

Data availability

All materials, stimulus lists, scripts, simulation and analysis code, and data from the experiments reported here can be found on the Open Science Framework (https://osf.io/g8q3n). Both Experiment 1 and 2 were preregistered on the Open Science Framework (Exp 1: https://osf.io/kmur8, Exp 2: https://osf.io/9qx78).

General model structure and approach

The basic architecture of the model consisted of a network of nodes representing linguistic information, based on Dell (1986). The model had four layers (see Figure 1): (a) a semantic layer, containing the semantic features of words (i.e., 10 unique nodes per word), (b) a lexical layer, in which each node represents a word (i.e., *sun*, *son*, *rat*, and *bed*), (c) a phonological layer, representing the individual phonemes of each word, and d) an orthographic layer with the individual graphemes (i.e., letters) of each word. A fifth layer, representing syntax (i.e., noun or verb), was only implemented in the third set of simulations. As mentioned earlier, some researchers posit separate lemmas, but one shared phonological lexeme for homophones (e.g., Biedermann et al., 2002; Jescheniak & Levelt, 1994), whereas others claim separate lexemes (e.g., Bonin & Fayol, 2002; Caramazza et al., 2001; Caramazza & Miozzo, 1998; Jacobs et al., 2004; Miozzo & Caramazza, 2005). However, all of them assume separate *orthographic* lexemes (since the spelling is different). For simplicity, we implement lexical items in a single layer[2], representing orthographic lexemes (or alternatively lemmas), which can be reconciled with both views.

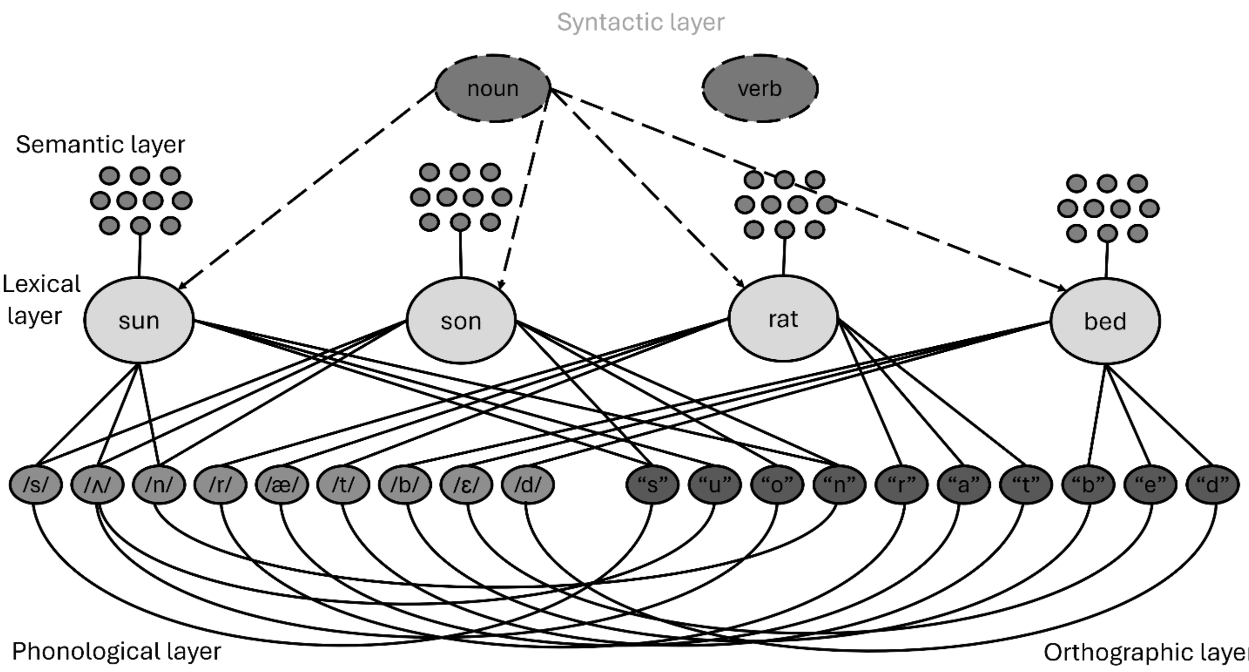

Figure 1. Architecture of the homophone model. The homophone target is "sun", while "son" acts as a competitor. The non-homophone target is "rat" with "bed" as competitor. The syntactic layer is only implemented in some of the simulations (see text).

[2] A difference between lemmas and lexemes could have meaningful consequences for neuropsychological studies, since the two types of representations may be lesioned separately. Since the current study targets neurotypical individuals, even if we modeled lemmas and lexemes as two separate layers, their activation would be highly correlated with one another, hence the choice of modeling them as one layer. For similar reasons, we do not use separate input and output layers for phonological and orthographic representations.

Note that we did not include semantic competitors in the lexical layer, because homophones are generally not semantically related to one another. Therefore, the inclusion of a semantically related competitor is not expected to affect homophone production any differently than non-homophone production. Similarly, we did not include phonological neighbors other than the homophone competitor, because non-homophone phonological competitors are also expected to affect homophone and non-homophone target words the same way. In short, the goal of the modeling here is not to fit quantitative error data or to estimate the probability of different kinds of errors on homophone targets, for which a larger set of lexical units that is more representative of the mental lexicon would be required. Rather, we set up the simulations to closely mirror the empirical design, with a homophone target (*sun*), a homophone competitor (*son*), a matched non-homophone target (*rat*), and a non-homophone competitor (*bed*).

Each word in the lexical layer was connected to its corresponding nodes in all other layers, and these connections were bidirectional (Nozari & Dell, 2009; Pinet & Nozari, 2018). In addition, there was a direct connection between the phonological layer and the orthographic layer through sound-to-spelling mapping rules (e.g., the sound /s/ was connected to the letter *s*). Since we controlled for lexical frequency in both experiments, frequency effects were not modeled in the current simulations. We also clamped the input to the model to simulate holding the phrase in verbal working memory as it was being typed. This means that the target phonological nodes keep receiving input across the different timesteps. The activation of an individual node was determined by Equation (1):

$$(1)\ A(j,t) = (A(j,t\text{-}1)\ (1\text{-q}) + \textstyle\sum_i w_{ij} \cdot A(i,t\text{-}1) + \sum_i w_{ij}\, fX_i)\ (1 + N(0,\text{a})) + N(0,\text{b})$$

where $A(j,\ t)$ is the activation of node j at timestep t, q is the decay parameter, w_{ij} is the weight of the connection from node i to j. The weight depends on the type of connection: s for semantic-lexical connections, p for lexical-phonological connections, o for lexical-orthographic connections, and po for phonological-orthographic connections. Clamping is implemented as $\Sigma w_{ij} \cdot f X_i$, where X is the binary input phoneme vector (if $X_i = 1$, phoneme i is present) and f is the input phoneme strength (i.e., the input that was originally given to the phonemes). Two sources of noise, activation noise a and intrinsic noise b were sampled from a normal distribution with mean 0 and standard deviation a and b, respectively.

Similar to the two-step interactive activation model (Dell, 1986), there were two steps here: a Step 1 in which input (see below) activated the lexical items compatible with that input and the lexical node with the highest activation was selected, and a Step 2, in which the selected lexical node received a jolt of 100 activation units for segmental encoding (see also Dell & O'Seaghdha, 1991; Foygel & Dell, 2000). In both steps, activation spread through the network for eight timesteps. The values of the model parameters were based on previous studies (e.g., Gauvin & Hartsuiker, 2020; Nozari et al., 2011): *s, p* = 0.04, *q* = 0.6, noise parameters *a* = 0.01 and *b* = 0.16. Since previous models did not model orthography, the values of *o* and *po* were set to be the same as for *s* and *p*. For each simulation, we ran 10,000 iterations and stored the output activation of each lexical node after Step 1, and the output activation of each orthographic node after Step 2 into a data frame.

The outcome measure was conflict (Nozari et al., 2011), computed as $-ln(activation_{target} - activation_{competitor})$, which can be computed at any layer of the system. The closer the activation of two nodes, the higher the value of conflict. Previous simulations have shown that high conflict is associated with more errors (Nozari et al., 2011). Empirically too, conditions that increase conflict have been shown to make production more difficult (Burgess & Nozari, 2022; Hanley et al., 2016; Gauvin & Hartsuiker, 2020; Nozari et al., 2011, 2019; Pinet & Nozari, 2021). As such, we used conflict as an index of production difficulty to be tested using the empirical measures obtained from Experiments 1 and 2. In Simulation 1, conflict between the homophone target and its competitor was compared to that between the non-homophone target and its competitor. Higher conflict for homophones indexes greater competition and would predict greater production difficulty, especially a slowdown, even on correct trials. In Simulation 2, conflict was compared between the target homophone and its competitor in single- vs. double-homophone sentences. Higher conflict in the latter points to increased competition and predicts greater production difficulty in HH compared to CH-HC sentences. Finally, in Simulation 3, conflict between the homophone target and its competitor was compared for same- vs. different-category homophones for a range of syntactic input to determine the range of possible syntactic influences on homophone production.

One of the utilities of simulations is to study the level(s) at which conflict arises. Theoretically, conflict in the homophone interference effect could arise in the lexical or orthographic layer or both. Computing conflict at the lexical layer is straightforward; activation of the target is simply compared to that of the competitor. However, to obtain a measure that indexes orthographic conflict, one must be careful to exclude errors that arise in the first stage of processing. To understand this, imagine that competitor “son” has been mistakenly selected in place of the target “sun” in Step 1, and Step 2 is completed without a glitch. In such a case, competition will be low in the orthographic layer. This is neither surprising nor does it reveal something about pure orthographic conflict. The real question is whether there are cases where the correct target has been selected at the lexical layer, but high conflict in the orthographic layer has caused the homophone interference effect. To answer this question, one must exclude errors in the first stage of processing, because including such errors could mask true orthographic competition. Whenever appropriate, we report both sets of conflict measures for lexical and orthographic layers in Simulations 1-3.

Simulation results and discussion

Simulation 1: Testing competition as a possible origin of the basic homophone interference effect

This simulation tested the basic homophone effect, i.e., whether having a phonologically identical competitor increases conflict when typing a target word compared to when the target has no such competitor. To this end, we once modeled typing a sentence with a homophone target word *sun* and once with a non-homophone target *rat,* and measured conflict between each target word and its respective competitor *(son* and *bed,* respectively*)*. A sentence dictation task contains a prominent phonological component. Thus, 0.4 units of activation were sent to each of the phonological nodes of the targets (/s/, /ʌ/, /n/ for the homophone target or /r/, /æ/, /t/ for the non-homophone target). The value of input was chosen to generate error rates within the range observed in Bonin et al. (2001). Given that sentential context naturally activates meaning (e.g., Halle & Stevens, 1962; ten Oever et al., 2022), we sent 10 units of activation to the ten semantic features of the target (*sun* or *rat*) as in the original two-step interactive model. Note that the input to semantics is much higher than input to phonology, because the assumption is that typing in neurotypical individuals is still primarily guided by meaning, even if the input is phonological. After eight timesteps (Step 1), we measured the conflict between the homophone target *sun* and its competitor *son* in one case, and between the non-homophone target *rat* and its unrelated competitor *bed* in the other case. After an additional eight timesteps (Step 2), we measured the conflict in the grapheme layer (after excluding lexical misselections, as we described above) for the vowels of the homophones (*u* and *o*) and non-homophones (*a* and *e*), as this is where the critical competition takes place. Higher conflict for the homophone target representations than the non-homophone target representations indexes the homophone interference effect. Moreover, the locus of conflict allows us to determine the level of representation at which the interference arises.

Results of Simulation 1 can be seen in Figure 2A. Step 1 showed higher conflict between the homophone target and its competitor (dark blue bar; *sun* vs. *son*: M = 2.58) than between the non-homophone target and its competitor (light blue bar; *rat* vs. *bed*: M = 1.80). Step 2 also showed higher conflict between the homophone target and its competitor (blue bar; *u* vs. *o*: $M = 2.56$) vs. the non-homophone target and its competitor (orange bar; *a* vs. *e*: $M = 1.63$).

Collectively, the results of Simulation 1 showed that competition, in the absence of any assumptions about poor spelling knowledge, is sufficient to produce the homophone interference effect. Furthermore, they showed that homophones induce competition at both lexical and orthographic layers. Recall that the conflict uncovered at the orthographic level cannot simply be considered a consequence of competition at the lexical level, since we deliberately selected trials with correct lexical selection to examine orthographic conflict. As such, Simulation 1 suggests a dual (lexical and orthographic) source of conflict underlying the homophone interference effect.

Simulation 2: The effect of lexical competition on homophone interference

In Simulation 2, we increased lexical competition by comparing the level of conflict on the target homophone *sun* and the competitor *son*, when the sentence either contained the homophone competitor *son* (the double-homophone or HH sentences in our experiments) or the non-homophone competitor *bed* instead (the single-homophone or CH/HC sentences in our experiments). Since the competitors are also heard in the sentence and their meaning is activated, we sent an additional jolt of activation of 1 unit to each of the 10 semantic nodes corresponding to the word *son* or *bed*. The rationale for the lower semantic jolt to the competitor is that the model is simulating the time point at which the target – and not the competitor – is being typed.

The results of Simulation 2 are shown in Figure 2B. For Step 1, lexical conflict was higher in the double-homophone condition ($M = 2.73$, dark green bar) than the single-homophone condition ($M = 2.61$, light green bar). However, for Step 2, there was no such difference (single: $M = 2.56$, light green bar; double: $M = 2.56$, dark green bar). Parameter space exploration showed that this pattern was not limited to the current values of jolts to the semantic, lexical, or orthographic nodes (Appendix A). These results predict that increasing lexical competition (e.g., in HH sentence), even when confined to the lexical layer, is expected to slightly increase the magnitude of the homophone interference effect compared to low lexical competition conditions (e.g., HC and CH sentences).

Simulation 3: The effect of syntax on homophone interference

Our previous two simulations took into account semantic, but not syntactic, input. Although processing meaning is arguably more automatic than syntactic structure, and the restrictive effect of syntax on lexical selection is even weaker in writing than in speaking (e.g., Hotopf, 1980; Romani et al., 2002), the latter certainly plays a role in production from meaning. This becomes particularly important in Experiment 2, when participants must answer questions. What is unclear is the extent to which syntactic input could override segmental interference. Simulation 3 addresses this question. For this simulation, we created a syntactic layer with two nodes (noun and verb), connected to corresponding lexical nodes to represent their syntactic category. We compared a situation in which the homophone target word *sun* and its competitor *son* belonged to the same syntactic category (i.e., noun) with a situation where the competitor *son* was artificially assigned to the verb category to mimic syntactic dissimilarity. The input to the model was the same as in Simulation 1 (i.e., 10 units of activation sent to each target semantic node and 0.4 units of activation to each phonological node), hence Step 1 was identical to that baseline model.

There are several ways to model syntactic input. In some models, syntax strictly limits lexical selection to the category cued by the syntactic node (e.g., only nouns; e.g., Dell, 1986). In other models, a syntactic cue excites the same-category lexical items and inhibits other-category lexical items (e.g., Gordon & Dell, 2003). These ways of implementing syntactic input over lexical selection naturally predict strong syntactic influences. While supported in spoken production, the evidence reviewed earlier is at odds with the predictions of these models in handwriting and typing. We, therefore, modeled syntactic input without restricting selection to the syntactic cue category or inhibitory effects of the syntactic cue on other-category items in order to examine the possible range of syntactic influences on lexical selection.

Syntactic input was modeled as part of the jolt after semantic-to-lexical mapping has been completed (Step 1).[3] This approach to implementing syntactic input has been proposed in previous versions of the two-step interactive model (e.g., Dell & O'Seaghdha, 1991). The rationale is that the primary factor in selecting a word is meaning, however, in sentence production, the jolt from the syntactic frame helps selecting a word from the syntactically appropriate category for that position in the sentence. For this reason, the syntactic input is modeled not as continuous spreading activation, but rather as a jolt at the point of lexical selection. But how strong should the syntactic input be? Unlike the phonological input parameter that can be tuned by examining error rates, the syntactic input is difficult to determine a priori. Moreover, since our main question is whether syntax has an influence on homophone interference or not, choosing any arbitrary value could bias the results. For these reasons, we opted to explore the full range of possibilities. We varied the syntactic jolt as a parameter ranging from 0 to 49% (just below half)

[3] We also implemented a version in which the syntax layer received input in Step 1 and spread activation, just like the other nodes in the network. This version yielded similar results with higher conflict for the same- vs. different-category homophones. The simulation code and output can be found on the OSF.

of the total 100-unit jolt. Larger values were not considered, as this would imply that syntactic input could override semantic input, which is implausible.

The syntactic part of the jolt was always sent to all the nodes compatible with the target's syntactic category. In the same-category condition, both the homophone target and the competitor received the jolt. In the different-category condition, only the homophone target received the jolt. The rest of the jolt was always given to the node with the highest lexical activation from the semantic features, as in other models. Figure 2C shows the results of Step 2 (since Step 1 is identical to Simulation 1).

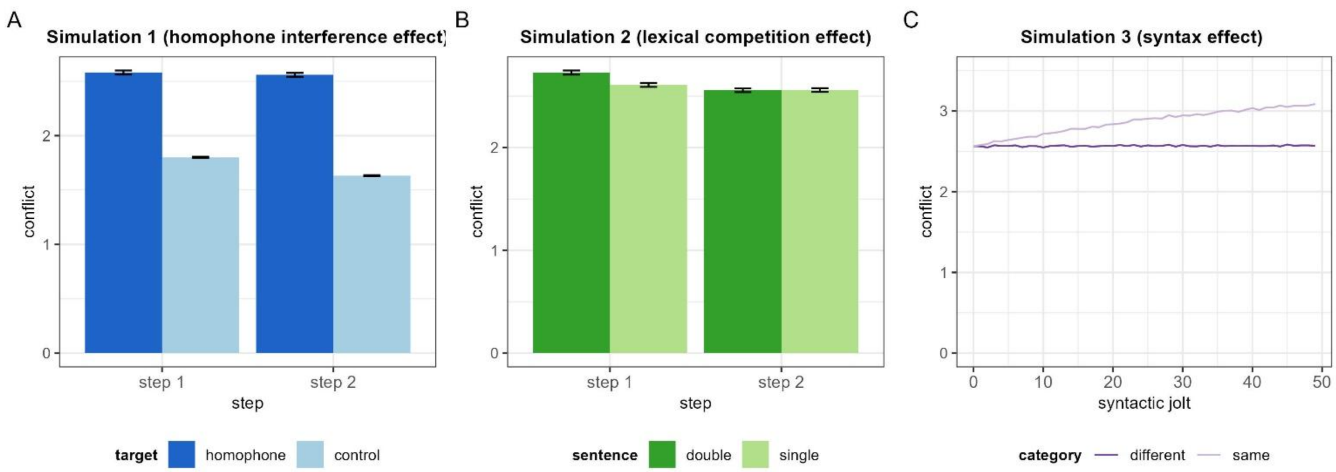

Figure 2. Results of Simulations 1-3. The dependent variable is conflict between the target and the competitor. Error bars represent 95% CI.

Simulation 3 (Figure 2C) showed a range of possible syntactic influences on the homophone interference effect. On the one extreme is very weak syntactic input, which unsurprisingly does not modulate the homophone interference effect. On the other extreme is the full-strength syntactic input, which causes greater increase in conflict in the same- vs. different-category homophones (0.52 or a 20% increase). Although this increase is modest compared to the basic homophone effect (0.93 or a 57% increase in Simulation 1), it shows that, in principle, syntax could modulate the homophone interference effect, if there is a strong syntactic input on lexical selection in typing (in line with White et al., 2010). In the following two experiments, we test where in this range the strength of syntactic input falls.

Experiment 1

Methods

Participants We collected data from 124 native speakers of American English (18-40 years old, M = 32.1, SD = 5.0; 89 males, 36 females), who were born in the US and lived there at the moment of testing. None of them reported any learning or language disorders. We recruited them through Prolific and paid them $5.60 for their participation. The sample size was determined by means of a power simulation using the mixedpower package in R (Kumle et al., 2021). If we assume a medium effect size (d = 0.50), 124 participants and 80 targets would yield a power of .86 to observe an interaction effect of lexical competition or syntax with homophone interference in onset latencies. Before being admitted to the study, candidates completed a typing prescreening test to ensure they were proficient typers (e.g., Pinet & Nozari, 2021, 2022). This study was conducted under the General Ethical Protocol of the Ethical Board Committee of the Faculty of Psychology and Educational Sciences at Ghent University and hence received automatic approval.

Materials We selected 40 English homophone word pairs, 20 from the same syntactic category (same-category pairs), and 20 from a different category (different-category pairs). To assess category membership, we used the part-of-speech information from the SUBTLEX-US database (Brysbaert et al., 2012). Pairs with potential dual use across two or more syntactic categories (e.g., 'brake'-'break', noun/verb) were not included to have the cleanest possible separation between same- and different-category pairs.

For each homophone, we selected a control word that was closely matched in terms of consonant-vowel structure (including gemination), word length, number of syllables, number of phonemes, and syntactic category. We also ensured that, on average, there was no difference in word frequency (Zipf-score, taken from SUBTLEX-US; van Heuven et al., 2014) and phonology-to-orthography (PO) consistency (Chee et al., 2020). For this final measure, we averaged the onset and rime PO token consistency values for each syllable of the target word[4] (see Appendix B). Furthermore, the matching

[4] We chose to use the token over the type consistency measure since it corrects for word frequency in the calculation.

controls always started with a letter that is typed with the same hand as the homophone in order to have no hand transition differences across conditions. Since we also compared same and different syntactic category homophone pairs, we ensured that, on average, stimuli in these conditions did not differ in terms of word frequency, orthographic Levenshtein distance (Levenshtein, 1966) between pairs, and PO consistency. We also ensured that the mean difference in Zipf-score between homophone mates was not significantly different across same (M= 1.35, SD= 1.08) and different syntax conditions (M = 1.21, SD = 0.73, t(38)= 0.45, p = .66).

The homophone pairs and their matched controls were combined into quadruplets with the sentences corresponding to the following conditions: homophone-homophone (HH), homophone-control (HC), control-homophone (CH), control-control (CC). For each homophone pair and their matched controls, we created a sentence context in which all combinations (HH, CH, HC, and CC) would result in a plausible sentence (see Appendix C.1 for the sentences). As such, there were 40 quadruplets of sentences appearing in four different conditions. We ensured that the average distance between the experimental words of interest, e.g., the two homophones or the homophone and control words, was similar across homophone pairs from the same vs. from a different category with a minimum of one word. The sentences were recorded by means of the AI-generated speech tool Descript (www.descript.com, using a female stock voice called 'Nancy'). In order to ensure that there was no difference in comprehensibility of sentences across conditions, we conducted a web-based norming study in which 40 native speakers of English judged the acceptability of spoken sentences. The 160 sentences were counterbalanced across four lists. Each participant judged 80 target sentences and 80 filler sentences that were not acceptable in terms of their semantics. Based on the norming study, we replaced five quadruplets containing low frequency homophones with more frequent alternatives. Reaction times analyses from the same norming study showed no significant difference across HH, CH, HC, and CC sentences, indicating that sentences were equally intelligible across conditions.

We created four lists of 40 sentences using a crossed design (i.e., from each quadruplet, one sentence appeared per list in a way that the conditions were counterbalanced across lists, without repeating two similar sentences within the same participant).

Procedure The experiment was programmed in JavaScript using the jsPsych library (de Leeuw, 2015) and hosted online via the JATOS platform (Lange et al., 2015) on a MindProbe server. After giving their informed consent, the participants completed the typing prescreening test, consisting of an untimed and a timed word typing task (15 words each). The goal of this task was to select touch-typists, i.e., those fluent enough in typing who could type without having to look at their hands (Pinet & Nozari, 2021). In order to pass the test, candidates had to reach >80% accuracy on the untimed task and <2000 ms total typing times for >80% of the words in the timed task (with >50% accuracy). Those that passed completed a short demographic form.

Next, participants were randomly assigned to one of the four lists in the sentence dictation task, in which they were asked to type the sentence they heard as fast and accurately as possible after hearing a beep sound at the end of the spoken sentence. They were prompted not to use any capitalization or punctuation. They then completed four practice trials. Each trial started with a fixation cross for 700 ms in the center of the screen. Next, the audio started playing while the response box appeared in the center of the screen. Participants could only listen to the sentence once. When the sentence was finished, a beep sound indicated that participants could initiate typing. Errors could be corrected as during usual typing, but all keypresses, including backspaces, were registered. Participants then pressed ENTER to continue with the next trial. There was no time constraint on typing the sentence. If participants started typing before the beep sound, a warning appeared prompting them to wait for the beep (only in practice trials). Finally, a blank screen appeared for 600 ms before continuing on to the next trial. There were 40 experimental trials in total, with a short break after 20 trials.

Finally, there was a spell-check task at the end. The purpose of this task was to probe the knowledge of correct spelling of homophones. This was implemented as a multiple-choice task in which participants were visually presented with the sentences from the dictation task, but with the target homophone left out. Participants were asked to complete the sentence using one of three options: the target, its actual homophone, and a misspelled homophone (e.g., *The fashion ____ turned out to be quite ugly*. Options: *idol*, *idle*, and **idel*). For sentences that appeared in the HH condition in the sentence dictation task, the same sentence appeared on two different trials with one of both targets left out and the other one replaced by the control word, in order to avoid influences from the competitor spelling being presented in the sentence. As such, this task also consisted of 40 trials. Once they finished, participants received the completion code and were paid.

Analyses We collected accuracy data, onset reaction times (RTs), and inter-key-intervals (IKIs) for all target words in the sentence dictation task. In order to capture interference effects before the onset of producing the target word, we also collected IKIs from the onset of the sentence up to the target word.[5] In

[5] This measure was not explicitly stated in the pre-registration, but we left open the possibility to use additional measures to study interference in the RTs.

case of technical problems or if a participant had a high number of incorrect responses in the sentence dictation task (i.e., > 50% of target responses), the person was excluded and replaced by a new participant (Experiment 1: *N*= 2; Experiment 2: *N*= 4). Since each sentence contained two target words, individual responses were scored once for the first and once for the second target. A target was excluded from the analyses if it was not correctly identified in the spell-check task.

There were four dependent variables (DVs): accuracy, onset RT, target duration, and pre-target duration. Responses to a target word were only scored as correct if no errors were made on the first typing attempt. Errors were divided into four categories (see Table 1 for examples): a) (pseudo)homophone errors, that were phonologically identical to the target, b) segmental errors, where a segment was dropped, misproduced or misplaced in such a way to change the word's pronunciation, c) lexical errors (when the target was replaced by another word or not produced at all), and d) morphosyntactic errors (e.g., plurals, different tenses). Lexical errors were excluded from the accuracy analysis, as participants did not attempt to produce the target word in these situations.

Onset RTs were operationalized as the time (in ms) needed to type the first letter of the target after pressing the spacebar after the previous word, whereas the target and pre-target duration (in ms) were calculated respectively by adding all individual IKIs of the target word and all individual IKIs before the target word, and dividing them by the number of characters. For the latency measures, we only included trials for which the responses were correct from the start of the sentence until the end of the target word on the first attempt, as errors and repair can slow down the typing process (e.g., Logan & Crump, 2010; Salthouse, 1986). Outlier raw RTs and mean IKIs (> 2.5 SDs above the group mean) were discarded from the analyses.

The data were analyzed using (generalized) linear mixed effects models by means of the lme4 (Bates et al., 2015) and afex packages (Singmann et al., 2016) in R. Factors were sum coded when involved in an interaction and otherwise dummy coded (to allow for a direct test of the effects without the need for post-hoc comparisons). For the random effects structure, we included Subject as subject-level and Quadruplet as item-level random effect. We always started from the full model, but when it did not converge, we reduced the random effects model following the recommendations by Matuschek et al. (2017). Three sets of models were tested corresponding to our three hypotheses[6]. All sets contained

[6] In the pre-registration, we stated that we would only test two sets of models, but since we have three research questions, it was more logical to have a separate model for the main effect of Word Type. Note that the results are similar with both approaches.

separate models with accuracy, onset RT, target duration, and pre-target duration as their DVs. Accordingly, the significance threshold was corrected to $\alpha = .0125$, using Bonferroni correction.

(1) Set 1. Is there a reliable homophone interference effect above and beyond the effects of orthographic consistency and knowledge of spelling? This set contained models with the following structure:

DV ~ Word Type + Zipf + (Word Type + Zipf | Subject) + (Word Type | Quadruplet)

Word Type was a factor with two levels (H vs. C) and (z-transformed) Zipf score was added as a control variable for frequency in the model.

(2) Set 2. Is there an effect of increased lexical competition? The models in this set only included homophone targets[7] and had the following structure:

DV ~ Competition + (Competition | Subject) + (1 | Quadruplet)

Here, Competition was a factor with two levels (single, double). 'Single' corresponds to homophones appearing in the HC or CH condition and 'double' to homophones appearing in the HH condition. Zipf score was not included in these models, since homophones are compared with themselves across conditions.

(3) Set 3. Is there an effect of syntactic category? Here, the model structure was defined as follows:

DV ~ Word Type*Category + Zipf + (Word Type*Category + Zipf | Subject) + (Word Type | Quadruplet)

Category was a factor with two levels (same vs. different category).

We also compared error types in the H vs. C condition to examine whether any homophone interference effect in accuracy is due to a higher proportion of (pseudo)homophone errors, and not simply any segmental error or typos, as in Bonin et al. (2001). For this comparison, we first performed a global χ^2 test of homogeneity on the distribution of errors, followed by pairwise proportion tests in the H vs. C condition (applying Bonferroni correction) with the focus on (pseudo)homophones and segmental errors.

Results

We excluded two homophone target items (124 observations or 2.5%) for which accuracy was < 50%. Also, 307 (6.2%) items that were selected incorrectly in the spell-check task were excluded. Furthermore,

[7] Originally, the idea was to make a comparison to control words, but since these are not of interest for our hypotheses regarding the lexical competition effect, we made a direct comparison between homophones in single vs. double homophone sentences.

136 (2.7%) lexical errors in the homophone condition and 235 (4.7%) in the control condition were not included in the main analyses.

Homophones vs. Controls Participants were less accurate when typing homophones vs. controls (β = -0.58, Z = -4.70, p < .001). The overall distribution of errors was significantly different across homophones and control words ($\chi^2(4)$ = 396.32, p < .001). Critically, there were more (pseudo)homophone errors in the H than in the C condition, whereas the proportion of segmental errors was similar across both conditions (see Table 1). As for the speed measures, there was no effect of Word Type in onset RTs (β = 3.05, t(42.5) = 0.66, p = .51) and target durations (β = 1.44, t(36.9) = 0.62, p = .54). However, pre-target durations were longer in homophones than in controls (β = 2.65, t(4316.9) = 3.70, p < .001; see Figure 3A). The model output for the fixed effects per DV can be found in Appendix D.1.

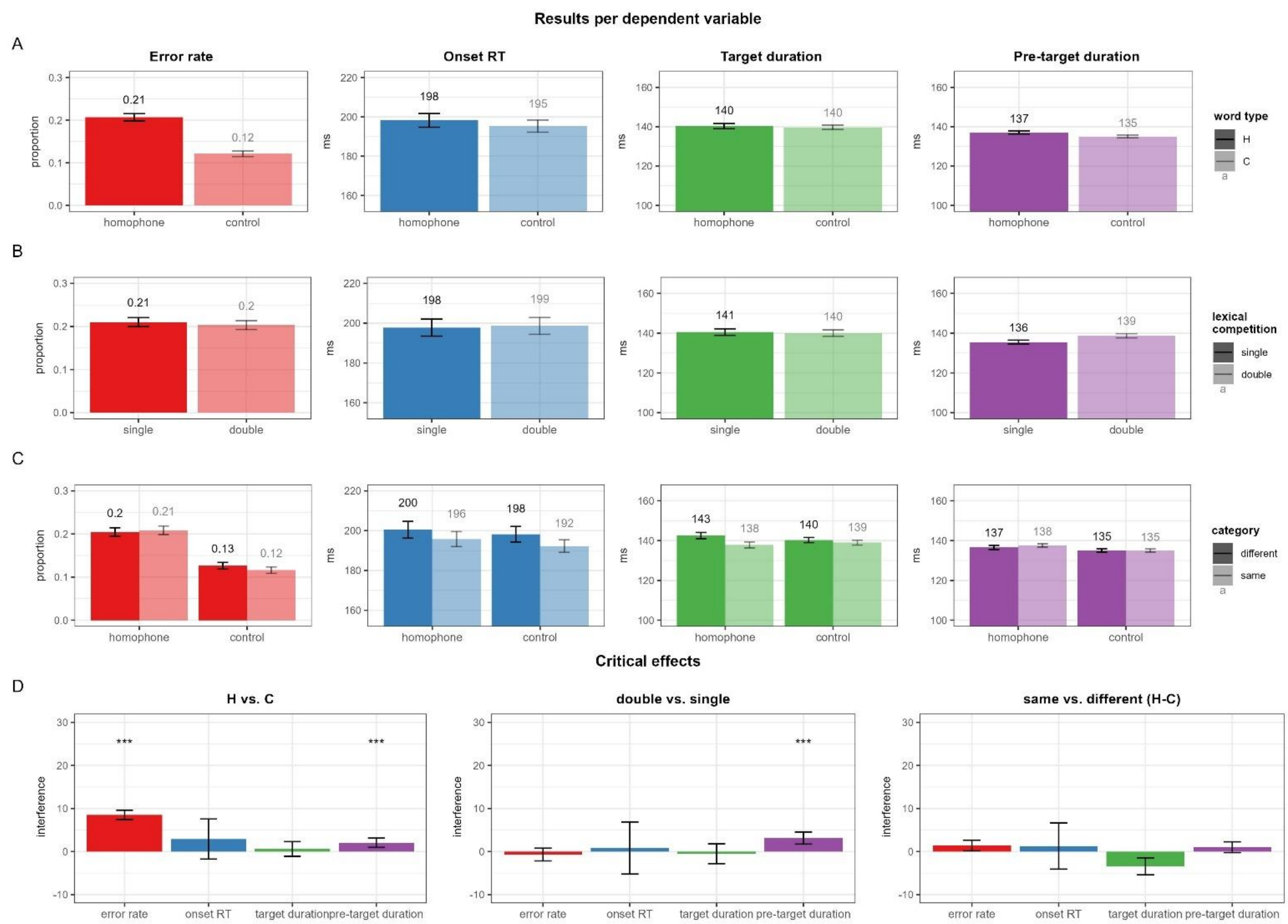

Figure 3. Results of Exp 1. Comparisons are shown for the four dependent variables (the four columns) for H vs. C (Panel A), single vs. double homophone (Panel B), and same vs. different category (Panel C). Panel D summarizes the interference effects for each row.

Table 1. Comparison of errors in Experiment 1, illustrated with examples for the target "idol". Percentages are relative to all responses, including correct ones.

Error type	Example	H (%)	C (%)	$\chi^2(1)$	*P*
(Pseudo)homophone	*idle/idel*	12.5*	2.2	378.40	< .001
Segmental	*odol/irdol*	6.5	7.7	5.01	.13
Lexical	*model/icon*	3.1	4.7		
Morpho-syntactic	*Idols*	0.9	1.6		

*55% of (pseudo)homophone errors were full homophone substitutions.

Single vs. Double Homophone There was no significant difference between single and double homophones in accuracy ($\beta = 0.06$, $Z = 0.76$, $p = .45$), onset RTs ($\beta = 3.33$, $t(2010.8) = 0.69$, $p = .49$), and target duration ($\beta = 0.04$, $t(102.5) = 0.02$, $p = .98$), but the pre-target duration was longer for the double compared to the single homophone condition ($\beta = 4.35$, $t(84.9) = 3.09$, $p < .01$; see Figure 3B and Appendix D.2).[8]

Same vs. Different Category The interaction between Word Type and Category was not significant in any DV (accuracy: $\beta = -0.02$, $Z = -0.25$, $p = .80$; onset RT: $\beta = 0.44$, $t(35.9) = 0.21$, $p = .84$; target duration: $\beta = -1.07$, $t(36.8) = -0.93$, $p = .36$; pre-target duration: $\beta = 0.12$, $t(4306.7) = 0.35$, $p = .73$), nor was there any main effect of Category (as assessed by type II Wald χ^2 tests, all *p*s > .50; see Figure 3C and Appendix D.3). Moreover, splitting up the same/different condition across double/single sentences yielded nearly identical accuracy rates across all homophone conditions.

Discussion

Experiment 1 showed that even after controlling for other factors in the design and ruling out spelling knowledge deficits, participants still committed more errors and had slower pre-target durations during the typing of homophones compared to control words. This finding matches the predictions of the competition account tested in Simulation 1. Moreover, as predicted by Simulation 2's results, increasing the lexical competition by having both homophones in the sentence led to slower pre-target durations compared to when there was only one homophone alternative in the sentence. These findings establish a

[8] Here and in Experiment 2: To minimize sublexical orthographic priming for the second target homophone, whenever possible, we chose controls that shared the critical segment with the homophone target. For example, 'bridle' and 'stable' both use the '-le' to represent the /əl/ sound. Since controlling for orthographic priming was not possible for all the materials, we also performed separate analyses on the first target homophone. These analyses yielded the same results as those collapsing target positions, showing that orthographic priming was not responsible for the reported effects.

true homophone interference effect rooted in competition. Interestingly, the homophone interference effect was not sensitive to syntactic category. Given Simulation 3's results, this finding corresponds to a situation of weak syntactic input, easily overwritten by other inputs. On the one hand, the absence of sensitivity to syntactic category matches anecdotal evidence on typing errors such as "there/their/they're". On the other hand, the finding is at odds with the report of White et al. (2010). As pointed out earlier, this difference likely stems from the absence of control sentences with non-homophone targets in White et al.'s experiment, without which it is difficult to rule out other differences between words belonging to different syntactic categories. Nevertheless, the discrepancy between the current and past results calls for a replication. Experiment 2 aimed to replicate all the findings of Experiment 1 but with a task that tapped more strongly into message construction from meaning.

Figure 4. Example picture from the question-answering task (for the idol-idle pair).

Experiment 2

Although spelling-to-dictation tasks are useful to elicit specific output in typing, they omit several key processes that are involved in everyday typing, including the transformation of ideas and concepts into messages, syntactic encoding, and lexical selection. To address this concern, in Experiment 2, we designed a question-answering task, in which participants were instructed to type replies to questions based on a picture while repeating all elements from that question. To that end, we slightly modified the stimuli from Experiment 1 to create questions (see 2a-2d):

(2a) HH: Who shouted at the fashion idol that he was quite idle?

(2b) HC: Who shouted at the fashion idol that he was quite ugly?

(2c) CH: Who shouted at the fashion icon that he was quite idle?

(2d) CC: Who shouted at the fashion icon that he was quite ugly?

(In this case, a picture of a shouting photographer was shown, see Figure 4, and participants typed *The photographer shouted at the fashion idol/icon that he was quite idle/ugly*). Although these questions were still orally presented, the participants had to process the meaning, retrieve new lexical items, and build a syntactic structure in order to answer them.

Methods

Participants A new sample of 124 participants took part in this experiment (age: M = 31.0, SD = 5.8; gender: 61 males, 62 females, 1 other). The selection criteria (including the pre-screening typing test) were the same as in the previous experiment. A new power simulation in which we updated the values of the parameters based on the model output of Experiment 1 yielded a power of .96 to detect an interaction effect at d= 0.50.

Materials We used the same target words and similar target sentences as in Experiment 1. Some of the targets were replaced if the proportion correct responses in the spell-check task was < .50 (N =1) or if the proportion correct final responses in the sentence dictation task was < .50 (N = 6). For the replacement, we chose homophone-pairs and controls that had a higher frequency than the original targets. As before, the mean difference in Zipf-score between homophone mates was not significantly different across same (M= 1.15, SD= 0.87) and different syntax conditions (M = 1.19, SD = 0.73, $t(38) = 0.18$, p = .85).

For each target sentence, we created a Wh-question that did not query the target words, but another aspect of the sentence (see Appendix C.2 for a list of sentences). As such, some target sentences were slightly altered, for instance by adding extra information (see 2a-2d). The Wh-questions were recorded in Descript, using the same voice as in Experiment 1. Next, we created a picture for each set of questions (N = 40) by means of an open-source AI text-to-image converter (dream.ai/create). We ensured that the same picture was appropriate for all versions of the same question. If possible, we avoided depicting the homophone words, since these could act as competitors in control sentences.

In order to check whether all questions and pictures were clear, we conducted a web-based norming study in which 40 native speakers of English were asked to answer the questions by using the picture. Based on the results, we adapted 18 pictures, mainly because there was low naming agreement (<.70 same response), and reformulated some of the questions to make the response more homogenous (e.g., *Where* was changed to *In which room*). Crucially, the participants gave a sensible response in at least 89% of the cases for each item, showing that the questions were clear and relatively easy to answer based on the picture.

Procedure The procedure was identical to Experiment 1, except that instead of hearing a sentence, participants heard a question, and typed a response based on a picture (this was illustrated with an example). They were instructed to respond as fast and as accurately as possible and repeat all elements from the question in their response without capitalization or punctuation. In order to practice the task, there were at least five practice trials, on which the subjects received feedback. Each trial started with a fixation cross that was presented for 700 ms in the center of the screen. Next, the audio started playing and once finished, the picture appeared in the center of the screen, with the response box below. Once participants finished typing, they pressed ENTER. In the practice phase, this was followed by a screen in which they saw the expected response appearing above their original responses, so they could compare. Finally, a blank screen appeared for 600 ms before continuing to the next trial. After five practice trials, participants continued onto the main task if at least 3/5 responses were correct. Otherwise, they got more blocks of five trials until they reached 3/5 accuracy in one block. There were 40 experimental trials divided into two blocks with a short break. Each block started with a catch trial, in which participants received feedback if their response was not correct in order to remind them to mention all elements from the question. These catch trials were not analyzed.

Analyses We ran the same analyses as in Experiment 1. Alpha was set to .0125 as in Experiment 1, to correct for multiple comparisons using the four measures.

Results

One homophone item was excluded due to accuracy below 50% (62 observations or 1.2%). We excluded another 371 (7.5%) observations for the homophones that had erroneous responses on the spell-check task and 302 (6%) lexical errors in the homophones and 439 (8.9%) of these errors in the controls.

Homophones vs. Controls Similar to Experiment 1, participants committed more errors in the homophones than in the control words (β = -0.36, Z = -4.33, p < .001), while the error pattern was significantly different across both word types ($\chi^2(4) = 320.91$, p < .001), with more (pseudo)homophone and fewer segmental errors in the homophones vs. controls (see Table 2). There was no difference between both word types in onset RTs (β = -0.15, $t(33)$ = -0.01, p = .98), target duration (β = 0.31, $t(37.9)$ = 0.08, p = .93), or pre-target duration (β = 2.36, $t(44.2)$ = 1.58, p = .12; see Figure 5A and Appendix E.1).

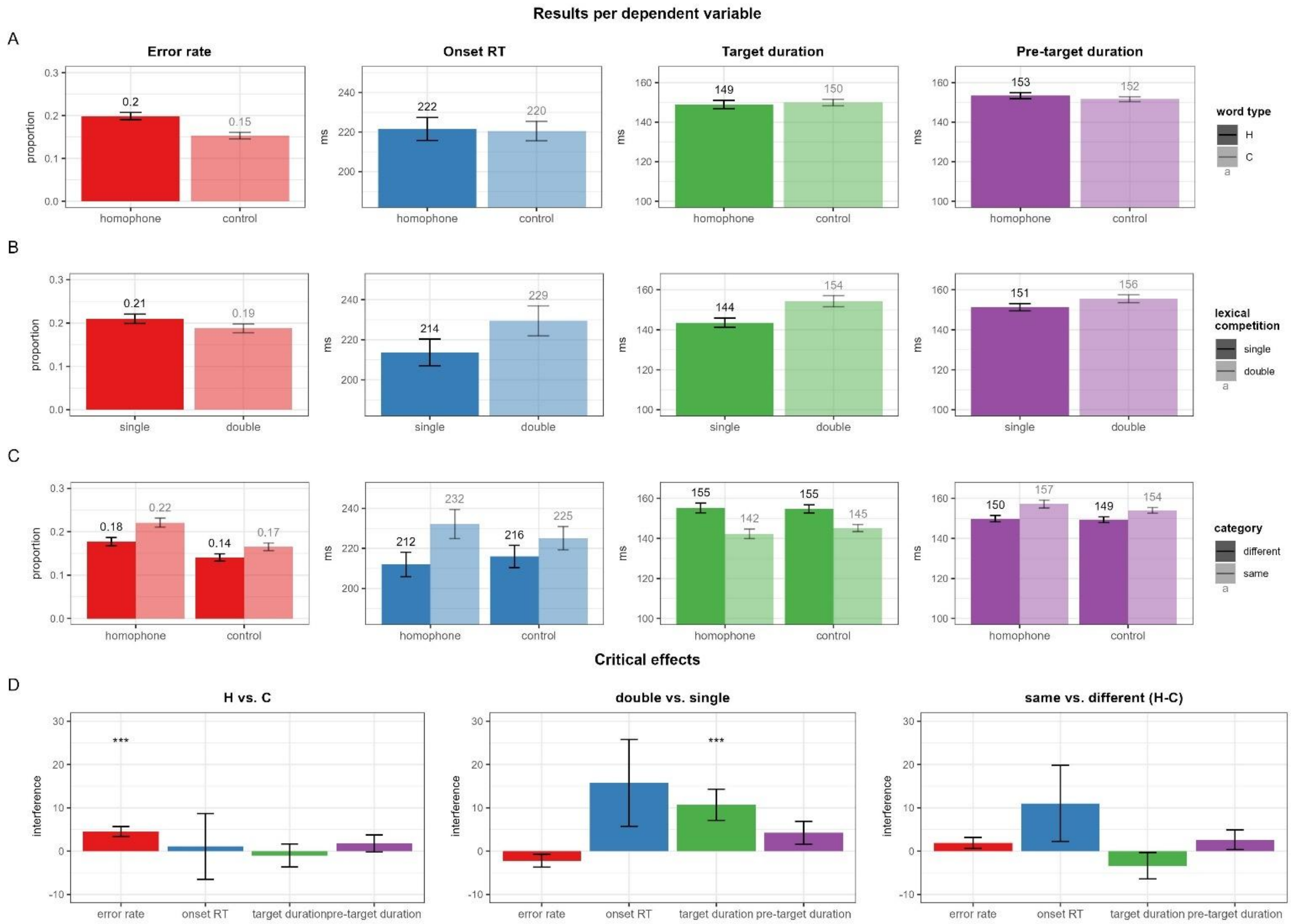

Figure 5. Results of Exp 2. Comparison of results across dependent variables for H vs. C (Panel A), single vs. double homophone (Panel B), and same vs. different category (Panel C) in Experiment 2. Panel D summarizes the interference effects.

Table 2. Comparison of errors in Experiment 2 illustrated with examples for the target "idol". Percentages are relative to all responses, including correct ones.

Error type	**Example**	**H (%)**	**C (%)**	**$\chi^2(1)$**	***P***
(Pseudo)homophone	*idle/idel*	9.7*	1.7	290.63	< .001
Segmental	*odol/isdol*	7.6	10.8	29.75	< .001
Lexical	*model/icon*	6.7	8.9		
Morpho-syntactic	*idols*	1.2	1.5		

*56% of (pseudo)homophone errors were full homophone substitutions.

Single vs. Double Homophone There was no significant difference in accuracy (β = 0.16, Z = 2.05, p = .04) and onset RTs (β = 18.69, $t(1256.2)$ = 2.20, p = .03) between the double and single homophone conditions, but participants had longer target durations in the double than in the single homophone condition (β = 10.40, t(1215.5) = 3.48, p < .001). Pre-target duration did not differ significantly across conditions (β = 4.97, t(982.0) = 2.20 p = .03; see Figure 5B and Appendix E.2).

Same vs. Different Category There was a main effect of Category in the target duration (Wald's type II $\chi^2(1) = 6.26$, p = .01), but not in the other measures (all ps > .09), which indicated that participants were slower in general when typing targets belonging to the same category sample compared to the different category sample. Critically, there was no Word Type-by-Category interaction effect in any DV (accuracy: β = -0.02, Z = -0.59, p = .56; onset RT: β = 1.80, t(33.9) = 0.47, p = .64; target duration: β = -1.41, t(7) = -0.78, p = .44; pre-target duration: β = 0.51, t(44.1) = 0.70, p = .49; see Figure 5C and Appendix E.3).

Table 3 summarizes the results across Experiments 1 and 2. Corrected alpha used for determining significant effects (marked by *) is .0125.

Table 3. Summary of effects across experiments, as expressed by p-values.

Exp 1	**Accuracy**	**Onset RT**	**Target duration**	**Pre-target duration**
Homophone interference effect	<.001*	.51	.54	<.001*
Lexical competition effect	.45	.49	.98	<.01*
Effect of syntax	.80	.84	.36	.73
Exp 2				
Homophone interference effect	<.001*	.98	.93	.12
Lexical competition effect	.04	.03	<.001*	.03
Effect of syntax	.56	.64	.44	.49

Discussion

Despite using a different task, Experiment 2 replicated the main result of Experiment 1, i.e., a significant homophone interference effect. The increase in errors for homophones vs. controls was again mainly stemming from pseudo(homophones). Also similar to Experiment 1, increasing lexical competition exaggerated the homophone interference effect, although the effect manifested as longer typing times for the homophone itself vs. the words leading up to the homophone as in Experiment 1. Both findings bolster the evidence in favor of competition as the source of homophone interference. Finally, we replicated the weak influence of syntax on resolving lexical competition in typing.

General Discussion

In this study, we investigated competition as a core mechanism underlying the homophone interference effect in typing. Since the competition account hinges on the dynamics of language production, we used the interactive two-step model of word production to generate testable predictions for this account. Our simulations predicted that (a) homophone interference should be observed, even in the absence of any assumption about poor spelling knowledge, and (b) the effect should grow larger by inducing greater lexical competition. Both predictions were supported and replicated across our two experiments. Both experiments elicited a robust homophone interference effect, reflected in higher error rates on homophones compared to controls, despite controlling for a host of potentially confounding variables and rejecting spelling-knowledge errors. Prediction (b) was also supported, as we observed a larger homophone interference effect under increased lexical competition, significant on durations before (Experiment 1) and on (Experiment 2) the target homophone, although trends were also evident on other measures in Experiment 2 (see Table 3). These results cement the role of competition in the greater difficulty associated with homophone production.

Our last simulation modeled the influence of syntax on lexical selection. We deliberately picked a model that did not strongly restrict selection based on syntactic category. The simulations showed a possible range of influence, leaving open the question of where in this spectrum the effect falls in empirical data. Neither experiment found a significant influence of syntax on the homophone interference effect. This implies that syntactic input to lexical selection in the orthographic modality is either weak or inconsistent across individuals, especially compared to the influence of phonology on orthography. Collectively, these results help unravel the mechanisms underlying lexical selection in typing as well as the specific nature of the representations involved. We will unpack each of these points below.

Competition as a core mechanism in typing

We found a robust, systematic homophone interference effect, naturally explained by competition (e.g., Morton, 1980; White et al., 2008, 2010, 2013). This effect aligns well with studies showing that form similarity causes competition in language production. Such interference is observed across various modalities of production, such as spoken, written, and typed production (Harrison et al., 2020; Pinet & Nozari, 2018, 2023; Rogers & Storkel, 1998; Sadat et al., 2014; Sullivan, 1999; Wheeldon, 2003), during both production of the known vocabulary and learning of new words (Breining et al., 2016; 2019; Waller et al., 2024), across both repetition as well as naming tasks (Sevald & Dell, 1994), and both within the same language and cross-linguistically as in cognate production (Martin & Nozari, 2021; Muscalu & Smiley, 2018; Muylle et al., 2022). Moreover, the homophone interference effect in production mirrors systematic processing difficulties in homophone comprehension (Coltheart et al., 1994; van Orden, 1987).

The present findings do not support lack of spelling knowledge as the only source of homophone errors, as put forward by Bonin et al. (2001). As mentioned in the Introduction, the main problem with that study was measuring spelling knowledge in a group of participants different from the ones who completed the homophone study, which is problematic given the individual differences in vocabulary and spelling knowledge (Bonin et al., 2013). This issue notwithstanding, there are other differences between Bonin et al.'s (2001) study and ours that may have contributed to the contrasting findings. First, French is more orthographically transparent than English, which means that the overall prevalence of homophones may be different in the two languages (Bonin et al., 2001; 2015). It is more common in English than French for the same letter to be pronounced differently given the lexical context (e.g., hand and wand). Less orthographic transparency generally promotes formal errors (Bonin et al., 2001), which may explain the greater prevalence of homophone errors in our study in English. Second, Bonin et al. (2001) used a picture naming task. There are two key differences between picture naming and the tasks used in this study. (a) Picture naming taps into word production in isolation, whereas our tasks tap into word production in context. (b) Picture naming has no auditory input, whereas our tasks do have auditory input. It is possible that homophone interference effects diminish in the absence of auditory input, since there is only top-down (conceptual or lexically driven), but no bottom-up (auditory driven) priming of the phonology. In line with this possibility, White et al. (2012) found that priming the meaning of the target homophone reduced homophone errors. Moreover, homophone interference effects in our study were less pronounced in the question-answering task (Exp 2) compared to the sentence dictation task (Exp 1). On the other hand, eliciting sentences rather than isolated words is more representative of everyday use of language in typed form. Correspondingly, the findings better match the common experience of typing homophone errors such as “there/their/they're”, which anecdotally arise during sentence typing. Thus, although we acknowledge that a lack of knowledge can explain part of the homophone errors that are observed in typing, we believe that Bonin et al. (2001) may have missed core competition effects that underlie lexical selection in typing in general and the homophone interference effect specifically, at least in languages like English.

Additional support for this claim came from an increase in the magnitude of the homophone interference effect when the sentence contained both homophones. At first glance, this finding contradicts the reports of facilitation of target production after the recent production of a homophone competitor (Jacobs et al., 2015). However, this facilitation was reported in *spoken* production. Since orthographic forms are not strongly activated during speaking (e.g., Alario et al., 2007; Chen et al., 2002; Roelofs, 2006), there is no reason to expect homophones to create interference in the spoken modality. Better aligned with our findings are reports of White et al. (2008, 2010, 2013), who found more homophone substitution errors when priming the distinctive spelling of the competitor during a sentence dictation task, both in handwriting and typing. Notably, both our priming manipulation and White and colleagues' ones had the same effect, even though our manipulation was primarily lexical (since target and competitor were phonologically identical and could elicit direct lexical substitution errors), whereas theirs was primarily segmental (the prime could not substitute the target directly, it could only prime segments that indirectly caused an error). The syntax manipulation of the primes in White et al. (2013) and the semantic manipulation in White et al. (2012) also tapped into lexical effects. This brings up the following question: what is the locus of competition in the homophone interference effect?

Our simulations of the basic homophone effect show that competition can arise at both lexical and orthographic levels. Lexical competition arises when the phonological string */sʌn/* activates both words *sun* and *son.* Orthographic competition arises when the phoneme */ʌ/* activates both segments [u] and [o]. Moreover, the interactivity within the system further enhances the mutual lexical and orthographic influences on one another. Once activated, both *sun* and son will send activation to their orthographical segments, while the orthographic layer sends back activation to them. To show this abstract idea more concretely, Simulation 1's results demonstrated increased competition at both lexical and orthographic levels in typing. While the idea of orthographic competition in homophone production is intuitive, lexical competition may be accepted more hesitantly. However, in Simulation 2 we demonstrated that adding a second homophone could increase homophone interference solely by increasing competition and conflict at the lexical layer.

To summarize, our findings support lexically influenced competition as the source of the homophone interference effect. This is important as it fits with the growing body of literature pointing to a similar kind of competition across spoken and typed modalities. An example of such competition is the *repeated letter* effect (Pinet & Nozari, 2018). Pinet and Nozari (2018) showed that the probability of a consonant error was higher when words shared a vowel, e.g., p (*tip fig* → *fip tig*) > p (*top fig* → *fop tig*). The reason is that activation of the segment [i] in *tip* activates the word *fig* through feedback, which then activates its segment [f] increasing its competition with [t]. When the vowel is different (e.g., *top/fig*), feedback does not activate the competing word, rendering the probability of competition at the segmental level lower. Interestingly, parallel effects have been reported in spoken production (e.g., the repeated phoneme effect; Dell, 1986), pointing to common processing principles in spoken and typed modalities.

These parallels are particularly interesting in light of the many differences between speaking and typing. First, there is the different age of acquisition for both skills: children typically start producing oral linguistic utterances within the first 18 months of their life, whereas typing is a skill that often comes much later in life (i.e., after the age of 7, when children learn to read and write). Because of the later onset, typing is often considered less automatic than speech and this may impact the way in which we prepare for production. Second, skilled typists produce on average about 30-40 words per minute (Rumelhart & Norman, 1982), whereas the average speech rate is about 200 words per minute (Yuan et al., 2006). Third, typing is characterized as a discrete process at the motor level (i.e., it involves sequential keystrokes), whereas this is much less the case for speaking due to coarticulation. Finally, when we type, we have visual feedback that is usually not restricted in time, which allows us to go back and correct the utterance. In contrast, feedback from speech is fleeting: the moment the word has been spoken, it only leaves a trace in our working memory, which makes it much harder to apply corrections. In keeping with this, Pinet and Nozari (2021) showed that removal of visual feedback severely deteriorated error correction in typing. Despite these differences, the similarities observed in many aspects of language processing across different modalities speak to domain-generality at the level of computational principles that govern production (Nozari & Martin, 2024; Nozari & Novick, 2017; see Nozari et al., 2025, for a review).

Syntactic influences in lexical selection in typing

In two experiments, we found no evidence for the effect of syntax on homophone production. This is in contrast to White et al.'s (2010) report. At first glance, one may attribute our null effects to a lack of statistical power. However, this explanation is unlikely to explain the discrepancy: White et al.'s study had a large effect size of 0.85. If this was the true effect size, we should have had a power of >0.99 to detect a significant effect. In fact, we selected our sample size conservatively, to have a power of 0.85 to detect a medium effect size of 0.5 and still found a null result across both experiments.

The discrepancy is likely due to the absence of control sentences with matched non-homophone targets in White et al.'s (2010) study. To illustrate this point, consider the accuracy rates in Figure 5C (left). When leaving out the control condition (right half of the graph), one may interpret the higher accuracy for different- vs. same-category homophones as evidence for the effect of syntax. However, since a similar increase in accuracy is observed in the control words, the difference across categories cannot be due to syntactic overlap, but must be due to other factors (e.g., more difficult sentences in the same-syntax condition). While we cannot further speculate on the source of such differences between conditions, it is clear that these cannot be attributed to our syntax manipulation. Although these findings do not categorically refute a possible effect of syntax on restricting lexical access in typing, they show that this influence is not robust at the population level. This conclusion fits well with often-observed syntactic violations in common homophone errors (e.g., the "they're/their/there" substitution), as well as past reports on such violations (e.g., Connors & Lunsford, 1992; Largy et al., 1996). The weak contribution of syntax to the homophone interference effect highlights an area of difference between spoken and typed language production, with ramifications for both the nature of homophone representations and, more generally, theories of typing. While syntax strongly restricts the choice of lexical items in spoken production, as demonstrated in speech error patterns (see Garrett, 1975, 1976), it clearly has a much less restrictive influence on lexical selection in typing. This, in turn, means that syntax is not represented at a modality-independent lemma level, but at a modality-specific lexeme level (e.g., Caramazza & Hillis, 1991; Starreveld & La Heij, 2004). In terms of homophone representations, this finding supports the double-lexeme account of homophones. In terms of general theories of lexical selection, this result shows that syntax cannot be implemented in models of typing the same way as it is implemented in models of speaking. Neither the strict restriction of lexical selection to the cued syntactic category (Dell, 1986) nor the substantial inhibition of the non-category lexical representations (e.g., Gordon & Dell, 2003) is compatible with the current findings. Instead, these results call for typing models in which syntactic cues are much less influential during lexical selection.

But why is lexical selection in typing (and handwriting) less susceptible to syntactic influences than in speaking? Likely because of the role that phonology plays in handwriting and typing (e.g., Afonso & Álvarez, 2011; Bonin et al., 2001, 2015; Damian et al., 2011; Delattre et al., 2006; Pinet et al., 2016; Pinet & Martin, 2024; Zhang & Damian, 2010). The role of phonology in writing has been disputed amongst writing researchers (see Tainturier & Rapp, 2001, for a review), with stances ranging from the idea that phonological access is a necessary step in retrieving spelling representations (i.e., *the obligatory phonological mediation hypothesis,* e.g., Afonso & Álvarez, 2011; Geschwind, 1969; Luria, 1970) to the idea that spelling representations can be retrieved directly from lexical representations, without the need to pass through phonology (i.e., *the orthographic autonomy hypothesis*, e.g., Bonin et al., 1998; Miceli et al., 1997; Rapp et al., 1997; Zhang & Wang, 2015). Homophone substitution errors and other phonologically plausible errors are often taken as evidence for phonological mediation (Aitchison & Todd, 1982), whereas neuropsychological studies observing a dissociation between speech and writing are often brought up in favor of orthographic autonomy (e.g., Rapp et al., 1997; Shelton & Weinrich, 1997).

The contrasting views and evidence can be explained by dual-route theories of writing (e.g., Barry, 1994; Bonin et al., 2015; Houghton & Zorzi, 2003; Purcell et al., 2011; Rapp et al., 2002). Although there may be some differences in the details, the core assumption of these theories is that spelling representations can be accessed via two routes during production: a *direct route* connecting the lexical representations to the spelling representations, and an *indirect route* via sound representations, where sounds are connected to spelling representations through phoneme-grapheme conversion rules (e.g., /k/ is connected to *c* and *k*). These theories assume that both routes contribute to spelling retrieval in healthy adults. The dual-route theory is supported by empirical evidence in both handwriting (e.g., Bonin et al., 2001, 2015; Damian et al., 2011; Delattre et al., 2006; Zhang & Damian, 2010) and typing (Pinet et al., 2016; Pinet & Martin, 2024).

Research on the dual-route framework clearly shows that phonology plays a direct and important role in orthographic production. The question is whether phonological influences are strong enough to override syntactic influences on lexical selection. Our findings suggest that they are, at least in tasks with some auditory input. However, we acknowledge that the current study includes overt phonological input, which may tip the balance of phonological vs. syntactic input in favor of phonology. Under such circumstances, we can confidently conclude that syntax does not strongly influence lexical selection. But generalization of this claim to all situations requires testing the homophone effect within the context of tasks that provide no overt phonological input. This is a great avenue for future research.

Conclusions

The aim of the present study was to uncover the mechanisms underlying the homophone interference effect, and more generally, lexical selection in typing. We found converging support for competition during lexical selection in typing, giving rise to homophone interference. Furthermore, there was no clear modulating effect of syntax on homophone interference, indicating a less prominent role of syntax in lexical selection during typing as opposed to speaking.

Acknowledgements

This work was supported by a grant from Ghent University Special Research Fund to M.M. and the NSF BCS-2317121 grant to N.N.

References

Afonso, O., & Álvarez, C. J. (2011). Phonological effects in handwriting production: Evidence from the implicit priming paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(6), 1474–1483. https://doi.org/10.1037/a0024515

Aitchison, J., & Todd, P. (1982). Slips of the mind and slips of the pen. In R. N. St. Clair & W. von Raffler-Engel (Eds.), *Language and cognitive styles: Patterns of neurolinguistic and psycholinguistic development* (pp. 180–194). Swets & Zeitlinger.

Alario, F. X., Perre, L., Castel, C., & Ziegler, J. C. (2007). The role of orthography in speech production revisited. *Cognition*, *102*(3), 464-475. https://doi.org/10.1016/j.cognition.2006.02.002

Bailey, L. M., Lockary, K., & Higby, E. (2024). Cross-linguistic influence in the bilingual lexicon: Evidence for ubiquitous facilitation and context-dependent interference effects on lexical processing. *Bilingualism: Language and Cognition*, *27*(3), 495–514. https://doi.org/10.1017/S1366728923000597

Barry, C. (1994). Spelling routes (or roots or rutes). In G.D.A. Brown and N.C. Ellis (eds.), *Handbook of spelling: Theory, process and intervention* (pp. 27–49). Chichester, UK: John Wiley.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Biedermann, B., Blanken, G., & Nickels, L. (2002). The representation of homophones: Evidence from remediation. *Aphasiology*, *16*(10–11), 1115–1136. https://doi.org/10.1080/02687030244000545

Biedermann, B., Coltheart, M., Nickels, L., & Saunders, S. (2009). Effects of homophony on reading aloud: Implications for models of speech production. *Language and Cognitive Processes*, *24*(6), 804–842. https://doi.org/10.1080/01690960802597243

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). London: Academic Press.

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology*, *23*(1), 45-93. https://doi.org/10.1016/0010-0285(91)90003-7

Bonin, P., & Fayol, M. (2002). Frequency effects in the written and spoken production of homophonic picture names. *European Journal of Cognitive Psychology*, *14*(3), 289–313. https://doi.org/10.1080/09541440143000078

Bonin, P., Fayol, M., & Peereman, R. (1998). Masked form priming in writing words from pictures: Evidence for direct retrieval of orthographic codes. *Acta Psychologica*, *99*(3), 311–328. https://doi.org/10.1016/S0001-6918(98)00017-1

Bonin, P., Méot, A., Lagarrigue, A., & Roux, S. (2015). Written object naming, spelling to dictation, and immediate copying: Different tasks, different pathways? *Quarterly Journal of Experimental Psychology*, *68*(7), 1268–1294. https://doi.org/10.1080/17470218.2014.978877

Bonin, P., Méot, A., Millotte, S., & Barry, C. (2013). Individual differences in adult handwritten spelling-to-dictation. *Frontiers in Psychology, 4,* 1–11.

Bonin, P., Peereman, R., & Fayol, M. (2001). Do phonological codes constrain the selection of orthographic codes in written picture naming? *Journal of Memory and Language*, *45*(4), 688–720. https://doi.org/10.1006/jmla.2000.2786

Breining, B., Nozari, N., & Rapp, B. (2016). Does segmental overlap help or hurt? Evidence from blocked cyclic naming in spoken and written production. *Psychonomic Bulletin & Review*, *23*(2), 500–506. https://doi.org/10.3758/s13423-015-0900-x

Breining, B. L., Nozari, N., & Rapp, B. (2019). Learning in complex, multi-component cognitive systems: Different learning challenges within the same system. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 45*(6), 1093–1106. https://doi.org/10.1037/xlm0000630

Brysbaert, M., New, B., & Keuleers, E. (2012). Adding part-of-speech information to the SUBTLEX-US word frequencies. *Behavior Research Methods*, *44*(4), 991–997. https://doi.org/10.3758/s13428-012-0190-4

Burgess, J., & Nozari, N. (2022). A conflict-based model of speech error repairs in humans. In J. Culbertson, A. Perfors, H. Rabagliati, & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society.* Cognitive Science Society.

Caramazza, A., Costa, A., Miozzo, M., & Bi, Y. (2001). The specific-word frequency effect: Implications for the representation of homophones in speech production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*(6), 1430–1450. https://doi.org/10.1037/0278-7393.27.6.1430

Caramazza, A., & Hillis, A. E. (1991). Lexical organization of nouns and verbs in the brain. *Nature*, *349*(6312), 788-790. https://doi.org/10.1038/349788a0

Caramazza, A., & Miozzo, M. (1998). More is not always better: A response to Roelofs, Meyer, and Levelt. *Cognition*, *69*(2), 231–241. https://doi.org/10.1016/S0010-0277(98)00057-2

Chee, Q. W., Chow, K. J., Yap, M. J., & Goh, W. D. (2020). Consistency norms for 37,677 English words. *Behavior Research Methods*, *52*(6), 2535–2555. https://doi.org/10.3758/s13428-020-01391-7

Chen, J.-Y., Chen, T.-M., & Dell, G. S. (2002). Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language, 46*(4), 751–781. https://doi.org/10.1006/jmla.2001.2825

Coltheart, V., Patterson, K., & Leahy, J. (1994). When a ROWS is a ROSE: Phonological Effects in Written Word Comprehension. *The Quarterly Journal of Experimental Psychology Section A*, *47*(4), 917–955. https://doi.org/10.1080/14640749408401102

Connors, R. J., & Lunsford, A. A. (1992). Exorcising Demonolatry. *Written Communication*, *9*(3), 404–428. https://doi.org/10.1177/0741088392009003004

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

Damian, M. F., Dorjee, D., & Stadthagen-Gonzalez, H. (2011). Long-term repetition priming in spoken and written word production: Evidence for a contribution of phonology to handwriting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 813–826. https://doi.org/10.1037/a0023260

Damian, M. F., Vigliocco, G., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, *81*(3), B77–B86. https://doi.org/10.1016/S0010-0277(01)00135-4

Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Sound-to-spelling regularity affects both writing latencies and durations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6), 1330. https://doi.org/10.1037/0278-7393.32.6.1330

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Dell, G. S. (1984). Representation of serial order in speech: evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*(2), 222. https://doi.org/10.1037/0278-7393.10.2.222

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*. 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

Ferreira, V. S., & Slevc, L. R. (2007). Grammatical encoding. *The Oxford handbook of psycholinguistics*, 453-469.

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

Frith, U. (1980). Unexpected spelling problems. In U. Frith (Ed.), *Cognitive processes in spelling*, pp. 495–515. London: Academic Press.

Garrett, M. F. (1975). The analysis of sentence production. In *Psychology of learning and motivation* (Vol. 9, pp. 133-177). Academic Press. https://doi.org/10.1016/S0079-7421(08)60270-4

Garrett, M.F. (1976). Syntactic processes in sentence production. In R.J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 231–256). Amsterdam: North-Holland.

Gauvin, H. S., & Hartsuiker, R. J. (2020). Towards a new model of verbal monitoring. *Journal of Cognition*, *3*(1). https://doi.org/10.5334/joc.81

Geschwind, N. (1969). Problems in the anatomical understanding of the aphasias. In *Brain & Behavior* (1st edition, pp. 107–128). Routledge. https://doi.org/10.4324/9781315082035-5

Goldrick, M., Folk, J. R., & Rapp, B. (2010). Mrs. Malaprop's neighborhood: Using word errors to reveal neighborhood structure. *Journal of Memory and Language*, *62*(2), 113–134. https://doi.org/10.1016/j.jml.2009.11.008

Gordon, J. K., & Dell, G. S. (2003). Learning to divide the labor: An account of deficits in light and heavy verb production. *Cognitive Science*, *27*(1), 1-40. https://doi.org/10.1207/s15516709cog2701_1

Halle, M., & Stevens, K. (1962). Speech recognition: A model and a program for research. *IRE transactions on information theory*, *8*(2), 155-159. https://doi.org/10.1109/TIT.1962.1057686

Hanley, J. R., Cortis, C., Budd, M.-J., & Nozari, N. (2016). Did I say dog or cat? A study of semantic error detection and correction in children. *Journal of Experimental Child Psychology, 142,* 36–47. https://doi.org/10.1016/j.jecp.2015.09.008

Harrison, W., Hepner, C. R., & Nozari, N. (2020). Is segmental interference position-dependent? In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society. (pp. 681-687).* Cognitive Science Society.

Hartsuiker, R. J., Antón-Méndez, I., van Zee, M. (2001). Object attraction in subject-verb agreement construction. *Journal of Memory and Language*, *45*(4), 546-572. https://doi.org/10.1006/jmla.2000.2787

Hartsuiker, R. J., & Barkhuysen, P. N. (2006). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes*, *21*(1-3), 181-204.

Heller, J. R., & Goldrick, M. (2014). Grammatical constraints on phonological encoding in speech production. *Psychonomic Bulletin & Review, 21*(6), 1576–1582. https://doi.org/10.3758/s13423-014-0616-3

Hillis, A. E., & Caramazza, A. (1995). Representation of grammatical categories of words in the brain. *Journal of Cognitive Neuroscience*, *7*(3), 396-407. https://doi.org/10.1162/jocn.1995.7.3.396

Hotopf, W. H. N. (1980). Slips of the pen. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 287–307). Academic Press.

Hotopf, W. H. N. (1983). Lexical slips of the pen and the tongue: What they tell us about language production. In B. Butterworth (Ed.), *Language prodcution* (pp. 147–199). Academic Press.

Houghton, G., & Zorzi, M. (2003). Normal and impaired spelling in a connectionist dual-route architecture. In *Cognitive Neuropsychology*. https://doi.org/10.1080/02643290242000871

Jacobs, M., Singer, N., & Miozzo, M. (2004). The representation of homophones: Evidence from anomia. *Cognitive Neuropsychology*, *21*(8), 840–866. https://doi.org/10.1080/02643290342000573

Jacobs, C. L., Yiu, L. K., Watson, D. G., & Dell, G. S. (2015). Why are repeated words produced with reduced durations? Evidence from inner speech and homophone production. *Journal of Memory and Language*, *84*, 37-48. https://doi.org/10.1016/j.jml.2015.05.004

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

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

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, *33*(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kumle, L., Võ, M. L.-H., & Draschkow, D. (2021). Estimating power in (generalized) linear mixed models: An open introduction and tutorial in R. *Behavior Research Methods*, *53*(6), 2528–2543. https://doi.org/10.3758/s13428-021-01546-0

Lange, K., Kühn, S., & Filevich, E. (2015). "Just another tool for online studies" (JATOS): An easy solution for setup and management of web servers supporting online studies. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0130834

Largy, P. (1996). The Homophone Effect in Written French: The Case of Verb-Noun Inflection Errors. *Language and Cognitive Processes*, *11*(3), 217–256. https://doi.org/10.1080/016909696387178

Lastres López, C., & Manalastas, G. (2018). Errors in L1 and L2 university students' writing in English: Grammar, spelling and punctuation. *RAEL: Revista Electrónica de Lingüística Aplicada*, *16*(1), 118–134.

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. https://doi.org/10.1017/S0140525X99001776

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady*, *10*(8), 707–710.

Logan, G. D., & Crump, M. J. (2010). Cognitive illusions of authorship reveal hierarchical error detection in skilled typists. *Science (New York, N.Y.)*, *330*(6004), 683–686. https://doi.org/10.1126/science.1190483

Logan, G. D., & Crump, M. J. C. (2011). Hierarchical control of cognitive processes. In *Psychology of Learning and Motivation - Advances in Research and Theory* (pp. 1–27). https://doi.org/10.1016/B978-0-12-385527-5.00001-2

Luria, A. R. (1970). *Traumatic Aphasia*. De Gruyter Mouton. https://doi.org/10.1515/9783110816297

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, R. H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Mahon, B. Z., 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. https://doi.org/10.1037/0278-7393.33.3.503

Martin, C. D., & Nozari, N. (2021). Language control in bilingual production: Insights from error rate and error type in sentence production. *Bilingualism: Language and Cognition, 24(2), 374-388.* https://doi.org/10.1017/S1366728920000590

McCloskey, M., Macaruso, P., & Rapp, B. (2006). Grapheme-to-lexeme feedback in the spelling system: Evidence from a dysgraphic patient. *Cognitive Neuropsychology*, *23*(2), 278–307. https://doi.org/10.1080/02643290442000518

Miceli, G., Benvegnu, B., Capasso, R., & Caramazza, A. (1997). The independence of phonological and orthographic lexical forms: Evidence from aphasia. *Cognitive Neuropsychology*, *14*(1), 35–69. https://doi.org/10.1080/026432997381619

Miozzo, M., & Caramazza, A. (2005). The representation of homophones: Evidence from the distractor-frequency effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31* (6), 1360-1371. https://doi.org/10.1037/0278-7393.31.6.1360

Morton, J. (1980). The logogen model and orthographic structure. In U. Frith (Ed.), *Cognitive processes in spelling* (pp. 117–133). Academic Press.

Muscalu, L. M., & Smiley, P. A. (2019). The illusory benefit of cognates: Lexical facilitation followed by sublexical interference in a word typing task. *Bilingualism: Language and Cognition*, *22*(04), 848–865. https://doi.org/10.1017/S1366728918000792

Muylle, M., Van Assche, E., & Hartsuiker, R. J. (2022). Comparing the cognate effect in spoken and written second language word production. *Bilingualism: Language and Cognition*, *25*(1), 93–107. https://doi.org/10.1017/S1366728921000444

Navarrete, E., Del Prato, P., Peressotti, F., & Mahon, B. Z. (2014). Lexical selection is not by competition: Evidence from the blocked naming paradigm. *Journal of Memory and Language*, *76*, 253-272. https://doi.org/10.1016/j.jml.2014.05.003

Nozari, N., & Dell, G. S. (2009). More on lexical bias: How efficient can a "lexical editor" be? *Journal of Memory and Language*, *60*(2), 291–307. https://doi.org/10.1016/j.jml.2008.09.006

Nozari, N., Dell, G. S., & Schwartz, M. F. (2011). Is comprehension necessary for error detection? A conflict-based account of monitoring in speech production. *Cognitive Psychology*, *63*(1), 1–33. https://doi.org/10.1016/j.cogpsych.2011.05.001

Nozari, N., Freund, M., Breining, B., Rapp, B., & Gordon, B. (2016). Cognitive control during selection and repair in word production. *Language, Cognition and Neuroscience*, *31*(7), 886–903. https://doi.org/10.1080/23273798.2016.1157194

Nozari, N., & Hepner, C. R. (2019a). To select or to wait? The importance of criterion setting in debates of competitive lexical selection. *Cognitive Neuropsychology, 36*(5–6), 193–207. https://doi.org/10.1080/02643294.2018.1476335

Nozari, N., & Hepner, C. R. (2019b). To select or to wait? Response to the commentaries. *Cognitive Neuropsychology, 36*(5–6), 226–233. https://doi.org/10.1080/02643294.2019.1632280

Nozari, N., & Martin, R. C. (2024). Is working memory domain-general or domain-specific? *Trends in Cognitive Sciences*. Advance online publication. https://doi.org/10.1016/j.tics.2024.06.006

Nozari, N., Martin, C. D., & McCloskey, N. (2019). Is repairing speech errors an automatic or a controlled process? Insights from the relationship between error and repair probabilities in English and Spanish. *Language, Cognition and Neuroscience, 34*(9), 1230–1245. https://doi.org/10.1080/23273798.2019.1637007

Nozari, N., & Novick, J. (2017). Monitoring and control in language production. *Current Directions in Psychological Science, 26*(5), 403–410. https://doi.org/10.1177/0963721417702419

Nozari, N. & Omaki, A. (2018). Syntactic production is not independent of inhibitory control: Evidence from agreement attraction errors. In C. Kalish, M. Rau, J. Zhu, & T. Rogers (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 824-829).

Nozari, N. & Omaki, A. (2022). Revisiting Agreement: Do Children and Adults Compute Subject-verb Agreement Differently? In J. Culbertson, A. Perfors, H. Rabagliati, & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*. Cognitive Science Society. (pp. 409-415).

Nozari, N., & Pinet, S. (2020). A critical review of the behavioral, neuroimaging, and electrophysiological studies of co-activation of representations during word production. *Journal of Neurolinguistics, 53*. https://doi.org/10.1016/j.jneuroling.2019.100875

Nozari, N., Pinet, S., & Muylle, M. (2025). What can typing tell us about language production? *Annual Review of Linguistics, 12*. Article in press.

Oppenheim, G. M., Dell, G. S., & Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, *114*(2), 227-252. https://doi.org/10.1016/j.cognition.2009.09.007

Oppenheim, G. M., & Nozari, N. (2024). Similarity-induced interference or facilitation in language production reflects representation, not selection. *Cognition, 245, 105720.* https://doi.org/10.1016/j.cognition.2024.105720

Pinet, S., & Martin, C. D. (2024). Cross-modal interactions in language production: evidence from word learning. *Psychonomic bulletin & review*, 10.3758/s13423-024-02552-x. Advance online publication. https://doi.org/10.3758/s13423-024-02552-x

Pinet, S., & Nozari, N. (2018). "Twisting fingers": The case for interactivity in typed language production. *Psychonomic Bulletin & Review*, *25*(4), 1449–1457. https://doi.org/10.3758/s13423-018-1452-7

Pinet, S., & Nozari, N. (2021). The role of visual feedback in detecting and correcting typing errors: A signal detection approach. *Journal of Memory and Language*, *117*, 104193. https://doi.org/10.1016/j.jml.2020.104193

Pinet, S., & Nozari, N. (2022). Correction Without Consciousness in Complex Tasks: Evidence from Typing. *Journal of Cognition*, *5*(1). https://doi.org/10.5334/joc.202

Pinet, S. & Nozari, N. (2023). Different electrophysiological signatures of similarity-induced and Stroop-like interference in language production. *Journal of Cognitive Neuroscience, 6(35).* https://doi.org/10.1162/jocn_a_02014
Pinet, S., Ziegler, J. C., & Alario, F.-X. (2016). Typing is writing: Linguistic properties modulate typing execution. *Psychonomic Bulletin & Review*, *23*(6), 1898–1906. https://doi.org/10.3758/s13423-016-1044-3
Plaut, D. C., & Shallice, T. (1993). Perseverative and semantic influences on visual object naming errors in optic aphasia: A connectionist account. *Journal of Cognitive Neuroscience*, *5*(1), 89-117. https://doi.org/10.1162/jocn.1993.5.1.89
Purcell, J. J., Turkeltaub, P. E., Eden, G. F., & Rapp, B. (2011). Examining the central and peripheral processes of written word production through meta-analysis. *Frontiers in Psychology*, *2*, 239. https://doi.org/10.3389/fpsyg.2011.00239
Qu, Q., Damian, M. F., Zhang, Q., & Zhu, X. (2011). Phonology contributes to writing. *Psychological Science*, *22*(9), 1107–1112. https://doi.org/10.1177/0956797611417001
Rapp, B., Benzing, L., & Caramazza, A. (1997). The autonomy of lexical orthography. *Cognitive Neuropsychology*, *14*(1), 71–104. https://doi.org/10.1080/026432997381628
Rapp, B., & Caramazza, A. (1997). The modality-specific organization of grammatical categories: Evidence from impaired spoken and written sentence production. *Brain and Language*, *56*(2), 248-286. https://doi.org/10.1006/brln.1997.1735
Rapp, B., Epstein, C., & Tainturier, M.-J. (2002). The integration of information across lexical and sublexical processes in spelling. *Cognitive Neuropsychology*, *19*(1), 1–29. https://doi.org/10.1080/0264329014300060
Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review*, *107*(3), 460. https://doi.org/10.1037/0033-295X.107.3.460
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. (2006). The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychonomic Bulletin & Review*, *13*, 33-37. https://doi.org/10.3758/BF03193809
Roelofs, A. (2018). A unified computational account of cumulative semantic, semantic blocking, and semantic distractor effects in picture naming. *Cognition*, *172*, 59-72. https://doi.org/10.1016/j.cognition.2017.12.007
Rogers, M. A., & Storkel, H. L. (1998). Reprogramming phonologically similar utterances: The role of phonetic features in pre-motor encoding. *Journal of Speech, Language, and Hearing Research*, *41*(2), 258-274. https://doi.org/10.1044/jslhr.4102.258
Romani, C., Olson, A., Ward, J., & Ercolani, M. (2002). Formal lexical paragraphias in a single case study: how "masterpiece" can become "misterpieman" and "curiosity" "suretoy." *Brain and Language*, *83*(2), 300–334. https://doi.org/10.1016/S0093-934X(02)00022-6
Romani, C., Galluzzi, C., Bureca, I., & Olson, A. (2011). Effects of syllable structure in aphasic errors: Implications for a new model of speech production. *Cognitive psychology*, *62*(2), 151-192. https://doi.org/10.1016/j.cogpsych.2010.08.001
Rumelhart, D., & Norman, D. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science*, *6*(1), 1–36. https://doi.org/10.1016/S0364-0213(82)80004-9
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
Salthouse, T. A. (1986). Perceptual, cognitive, and motoric aspects of transcription typing. *Psychological Bulletin, 99*(3), 303-319. https://doi.org/10.1037/0033-2909.99.3.303
Scaltritti, M., Arfe, B., Torrance, M., & Peressotti, F. (2016). Typing pictures: Linguistic processing cascades into finger movements. *Cognition*, *156*, 16-29. https://doi.org/10.1016/j.cognition.2016.07.006
Schnur, T. T. (2014). The persistence of cumulative semantic interference during naming. *Journal of Memory and Language*, *75*, 27–44. https://doi.org/10.1016/j.jml.2014.04.006
Sevald, C. A., & Dell, G. S. (1994). The sequential cuing effect in speech production. *Cognition*, *53*(2), 91-127. https://doi.org/10.1016/0010-0277(94)90067-1
Shelton, J. R., & Weinrich, M. (1997). Further evidence of a dissociation between output phonological and orthographic lexicons: A case study. *Cognitive Neuropsychology*, *14*(1), 105–129. https://doi.org/10.1080/026432997381637
Singmann, H., Bolker, B., Westfall, J., Aust, F., Højsgaard, S., Fox, J., Lawrence, M. A., Mertens, U., & Love, J. (2016). afex: Analysis of factorial experiments. In *R package version 0.16-1. https://CRAN.R-project.org/package=afex*.
Starreveld, P., & La Heij, W. (2004). Phonological facilitation of grammatical gender retrieval. *Language and Cognitive Processes*, *19*(6), 677–711. https://doi.org/10.1080/01690960444000061
Stemberger, J. P. (1982). *The lexicon in a model of language production*. University of California, San Diego.
Sullivan, M. P. (1999). The nature of phonological encoding during spoken word retrieval. *Language and Cognitive Processes*, *14*(1), 15–45. https://doi.org/10.1080/016909699386365
Tainturier, M.-J., & Rapp, B. (2001). The spelling process. In B. Rapp (Ed.), *The Handbook of Cognitive Neuropsychology: What Deficits Reveal About the Human Mind* (1st ed.). Psychology Press.
Ten Oever, S., Carta, S., Kaufeld, G., & Martin, A. E. (2022). Neural tracking of phrases in spoken language comprehension is automatic and task-dependent. *Elife*, *11*, e77468. https://doi.org/10.7554/eLife.77468

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). Subtlex-UK: A New and Improved Word Frequency Database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition*, *15*(3), 181–198. https://doi.org/10.3758/BF03197716

Vigliocco, G., & Hartsuiker, R. J. (2002). The interplay of meaning, sound, and syntax in sentence production. *Psychological Bulletin*, *128*(3), 442–472. https://doi.org/10.1037/0033-2909.128.3.442

Waller, M., Yurovsky, D., & Nozari, N. (2024). Effect of similarity and training experiences on new vocabulary learning. In L. K. Samuelson, S. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.* Cognitive Science Society.

Wheeldon, L. (2003). Inhibitory form priming of spoken word production. *Language and Cognitive Processes*, *18*(1), 81–109. https://doi.org/10.1080/01690960143000470

White, K. K., Abrams, L., McWhite, C. B., & Hagler, H. L. (2010). Syntactic constraints in the retrieval of homophone orthography. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 160–169. https://doi.org/10.1037/a0017676

White, K. K., Abrams, L., Palm, G. M., & Protasi, M. A. (2012). Age-related influences on lexical selection and orthographic encoding during homophone spelling. *Psychology and Aging, 27*(1), 67–79. https://doi.org/10.1037/a0024422

White, K. K., Abrams, L., & Zoller, S. M. (2013). Perception-Production Asymmetries in Homophone Spelling: The Unique Influence of Aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *68*(5), 681–690. https://doi.org/10.1093/geronb/gbs099

White, K. K., Abrams, L., Zoller, S. M., & Gibson, S. M. (2008). Why did I right that? Factors that influence the production of homophone substitution errors. *Quarterly Journal of Experimental Psychology*, *61*(7), 977–985. https://doi.org/10.1080/17470210801943978

Yuan, J., Liberman, M., & Cieri, C. (2006). Towards an integrated understanding of speaking rate in conversation. In *Ninth International Conference on Spoken Language Processing*.

Zhang, Q., & Damian, M. F. (2010). Impact of phonology on the generation of handwritten responses: Evidence from picture-word interference tasks. *Memory & Cognition*, *38*(4), 519–528. https://doi.org/10.3758/MC.38.4.519

Zhang, Q., & Wang, C. (2015). Phonology is not accessed earlier than orthography in Chinese written production: evidence for the orthography autonomy hypothesis. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00448

Appendix

Appendix A: Explorations of the parameter space for Simulation 2

To find out whether the observed findings regarding the lexical competition manipulation (HH vs. HC/CH) depended on specific values for semantic, lexical, and orthographic input to the homophone competitor, we performed simulations with the following value ranges:

- semantic input: 0-5 (the maximum value of 5 was selected because above that, accuracy dropped below 70%, which does not match the empirical reports).
- lexical input: 0-0.30 (The maximum value 0.30 was chosen based on the mean activation of the target homophone's lexical node after eight timesteps in Step 1 for CH sentences, before the lexical boost was given).
- orthographic input: 0-0.35 (The maximum value 0.35 was selected to be lower than 0.4 phonological input).

We computed the difference in conflict between HH and CH to reflect the lexical competition effect. As the pattern did not change over the different values of lexical input, we picked four values for demonstration. The results of these explorations are visually presented in Figure A. Panel A, B, C, and D represent the explorations for lexical input being set to 0, 0.10, 0.20, and 0.30, respectively. The graphs on the left always show conflict at the lexical layer between the homophone target sun and the competitor son after Step 1 (before the lexical boost). The graphs on the right show the conflict at the grapheme layer between the homophone target letters "u" and the competitor letter "o" after Step 2. Conflict is shown as a function of semantic input (x axis) and orthographic input (y axis). Darker blue represents conflict values close to 0 (suggesting little to no lexical competition), while increasing values of conflict are shown in lighter blue.

As can be seen in Figure A, the left graphs show increased conflict with increasing values of the semantic input. This increase is gradual and near-monotonic across orthographic and lexical input values, with the latter two variables not substantially modulating the pattern. This pattern shows that semantic input is the predominant source creating lexical competition within the designated range of parameters (see above for the justification of this range). This finding naturally predicts no systematic effects in Step 2, which is supported by the pattern observed in the right graphs. In short, these explorations support the finding reported in the main text, namely that adding the homophone competitor to the sentence increases conflict in the lexical layer.

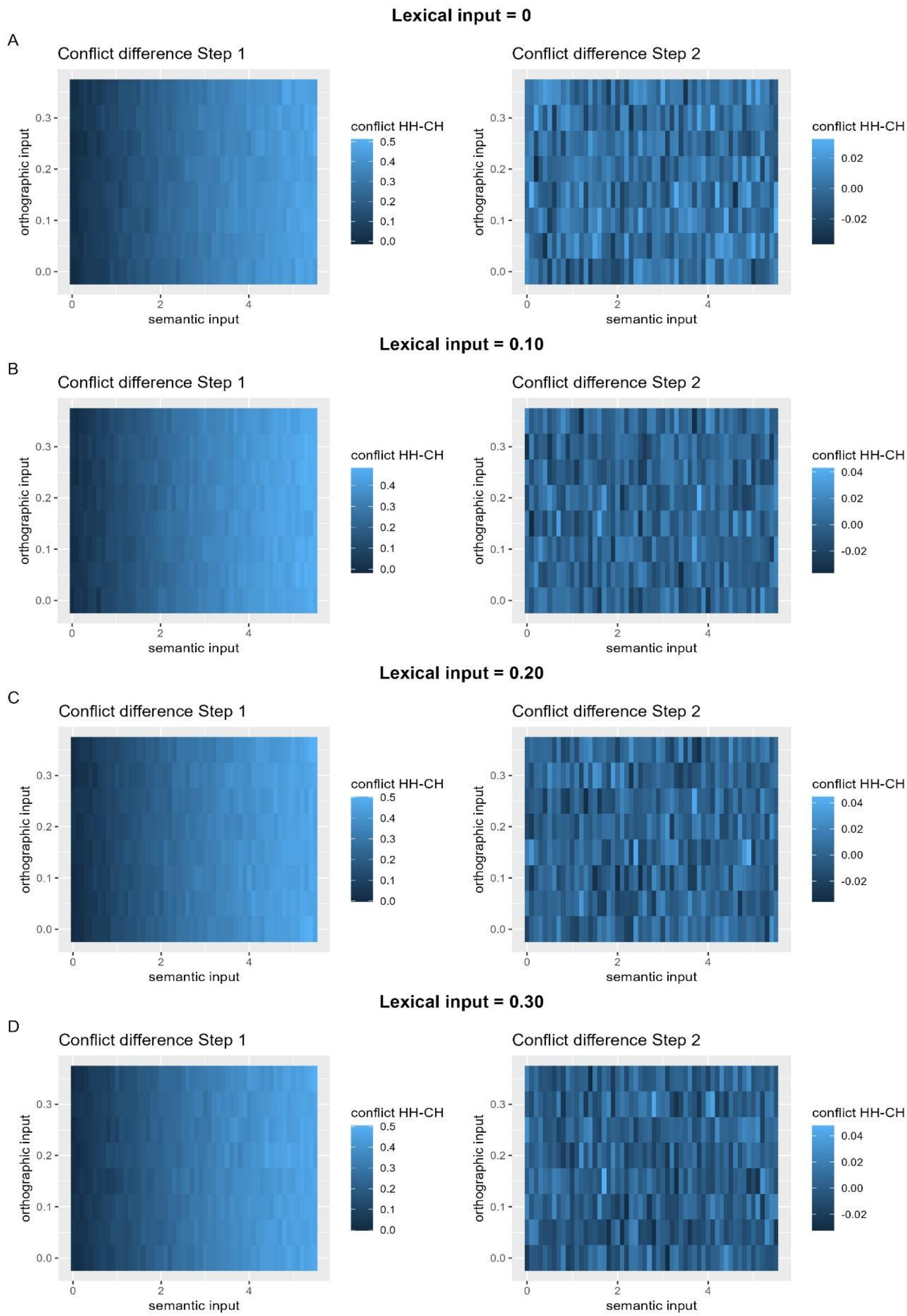

Figure A. Parameter explorations for the lexical competition effect (see text for explanation).

Appendix B. Characteristics of homophones and control words.

	Homophones *M (SD)*			**Control words *M (SD)***
	Same	Different	Total	
N letters	4.7 (1.0)	4.7 (1.4)	4.7 (1.2)	4.6 (1.1)
N syllables	1.2 (0.4)	1.3 (0.5)	1.2 (0.4)	1.2 (0.4)
N phonemes	3.6 (1.0)	3.6 (1.2)	3.6 (1.1)	3.7 (1.0)
Word frequency	4.1 (1.0)	4.2 (1.2)	4.1 (1.1)	4.3 (0.9)
PO consistency	.61 (.17)	.60 (.15)	.61 (.16)	.64 (.19)
N words between targets	2.9 (0.7)	2.8 (1.0)	2.9 (0.9)	2.9 (0.9)
OLD	1.8 (0.6)	2.3 (0.7)	2.1 (0.7)	4.1 (1.0)

Appendix C.1. Sentence materials used in Experiment 1

Sentence	Syntax condition
the donkey had a bridle/stable that was decorated with bridal/floral motifs	different
he guessed/fussed that the young guest/guard had arrived	different
we hear/mean they are coming here/home today	different
the holy/hazy mountains had become wholly/fully different	different
the fashion idol/icon turned out to be quite idle/ugly	different
it might/must be a mite/pine allergy	different
none/nine of them saw the nun/pup	different
the grandmother cleaned the rye/tie with a wry/shy smile	different
they seem/seek to repair the seam/rear of the dress	different
there/thus, we tend to follow their/clear rules	different
they threw/swung the ball through/toward the window	different
they were allowed/annoyed to speak aloud/again	different
the injured bald/wild man bawled/walked into the dark	different
the wealthy baron/rival had a barren/ragged estate	different
suddenly, I/he noticed an eye/ape staring at me	different
he received a loan/poem from the lone/late soldier	different
the marshal/master used to teach martial/manmade arts in the old days	different
the crowd massed/hopped around the mast/helm of the ship	different
he wore/tore his tunic in the war/bar	different
the drug addict searched in vain/fear for a usable vein/coin	different
if we find/risk it, we could get fined/fired	same
he crossed the gate/zone with a limping gait/goat	same
the moose/mouse tasted the lemon mousse/juice	same
eating British mussels/muffins strengthens the muscles/kingdom	same
he saw my son/dog in the sun/den	same
before playing the first chord/track, he attached a cord/belt to the guitar	same
the southern belle/babe rang the bell/boss	same
the seller/sinner left the hidden cellar/tunnel	same
the marine corps/force constitutes the core/fate of the camp	same
the guys/boys acted under the guise/voice of justice	same
he had his own manner/ladder to rebuild the manor/motel	same
the clock tolled/buzzed, as he told/wept about his misfortune	same
he grasped the escaping chipmunk and hare/mice by the hair/head	same
the young doe/cow smelled the fresh dough/roach	same
he felt a sharp pain/nail while moving the glass pane/pipe	same
he spilled some flour/drops on the red flower/flames	same
in the tale/cave, the hero struck the tail/beak of the dragon	same
he accomplished a major feat/goal with his feet/book	same
his beautiful piece/niece invited men to make peace/noise	same
the love of blue jeans/nails is in my genes/bones	same

Appendix C.2. Sentence materials used in Experiment 2

Question (answer)	Syntax condition
Which actor was seen/fed in the parking scene/space? (Brad Pitt)	different
Who guessed/fussed that the young guest/guard had arrived? (king)	different
How many dancers do we hear/mean are coming here/home today? (two)	different
Who noticed that the holy/hazy mountains had become wholly/fully different? (queen)	different
Who shouted at the fashion idol/icon that he was quite idle/ugly? (two)	different
What symptom confirmed that it might/must be a mite/pine allergy? (rash)	different
Who stated that none/nine of them had seen the nun/pup? (swimmers)	different
Which president passed/called the library in the past/last hour? (Joe Biden)	different
How many tailors seem/seek to repair the seam/rear of the dress? (two)	different
Where did the weak/deaf lawyer go for a week/scoop? (beach)	different
Who threw/swung the ball through/toward the window? (golfer)	different
How many kids were allowed/annoyed to speak aloud/again? (three)	different
Who made/left some lobster for the maid/jail? (chef)	different
Where did the baron/rival have a barren/ragged estate? (desert)	different
From where did I/he notice an eye/ape staring at me? (tree)	different
Who received a loan/poem from the lone/late soldier? (nurse)	different
What did they want to bury/carry near the berry/bunny farm? (piggy bank)	different
Who massed/hopped around the mast/helm of the ship? (pirates)	different
Who wore/tore his tunic in the war/bar? (pilot)	different
In which city did he search in vain/fear for a suitable vein/coin? (Paris)	different
Who warned us that if we find/risk it, we could get fined/fired? (judge)	same
Who crossed the gate/zone with a limping gait/goat? (cowboy)	same
With what did the moose/mouse taste the lemon mousse/juice? (straw)	same
Where did the man with the big muscles/pimples eat fresh mussels/muffins? (beach)	same
Which animal saw my son/dog in the sun/den? (fox)	same
To which instrument did he attach a cord/belt before playing the first chord/track? (guitar)	same
Where did the southern belle/babe ring the bell/boss? (pool)	same
With what did the seller/sinner leave the hidden cellar/tunnel? (bag)	same
Who spotted a fairy/daisy on the ferry/buggy? (baby)	same
Who judged that the guys/boys acted under the guise/voice of justice? (priest)	same
Who had his own manner/ladder to paint the manor/motel? (painter)	same
What did the knight/priest drink during the night/match? (wine)	same
In which room did he grasp the anxious hare/mice by the hair/head? (bathroom)	same
In which room did the young doe/cow smell the fresh dough/roast? (kitchen)	same
Who felt a sharp pain/nail while moving the glass pane/pipe? (doctor)	same
Which animal spilled some flour/drops on the beautiful flower/blazer? (raccoon)	same
Which animal's tail/beak did the hero strike in the tale/cave? (dragon)	same
Which president increased his stake/share in the steak/squid restaurant? (Donald Trump)	same
Whom did her beautiful piece/niece invite to make peace/noise? (soldier)	same
The love of what color jeans/nails is in my genes/bones? (blue)	same

Appendix D.1. Fixed effects of H vs. C models in Experiment 1 (sentence-dictation task)

Model formulas show random effects between brackets, with '‖' denoting uncorrelated random slopes.

A) Accuracy

Formula: Accuracy ~ Zipf + Word Type + (1 | Subject) + (Word Type | Quadruplet)

N= 9116

	Estimate	***SE***	***Z***	***p****
(Intercept)	2.15	0.10	21.41	<.001
Zipf	0.44	0.04	10.24	<.001
Word Type H	-0.58	0.12	-4.70	<.001

*alpha = 0.0125

B) Onset RT

Formula: Onset RT ~ Zipf + Word Type + (Zipf + Word Type | Subject) + (Word Type | Quadruplet)

N= 4667

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	201.40	8.55	98.77	23.56	<.001
Zipf	4.39	2.19	201.63	2.00	.047
Word Type H	3.05	4.61	42.47	0.66	.511

*alpha = 0.0125

C) Target Duration

Formula: Target Duration ~ Zipf + Word Type + (Zipf | Subject) + (Word Type | Quadruplet)

N= 4716

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	145.15	3.65	145.24	39.77	<.001
Zipf	-5.81	0.90	202.74	-6.48	<.001
Word Type H	1.45	2.34	36.90	0.62	.540

*alpha = 0.0125

D) Pre-target Duration

Formula: Pre-target Duration ~ Zipf + Word Type + (Zipf | Subject) + (1 | Quadruplet)

N= 4484

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	140.80	3.56	153.88	39.60	<.001
Zipf	-0.40	0.48	207.62	-0.83	.408
Word Type H	2.65	0.72	4316.89	3.70	<.001

*alpha = 0.0125

Appendix D.2. Fixed effects of lexical competition models in Experiment 1 (sentence-dictation task)

Model formulas show random effects between brackets, with ‘‖’ denoting uncorrelated random slopes.

A) Accuracy

Formula: Accuracy ~ Competition + (1 | Subject) + (1 | Quadruplet)

N= 4391

	Estimate	***SE***	***Z***	***p****
(Intercept)	1.48	0.14	10.91	<.001
Competition Double	0.06	0.08	0.76	.446

*alpha = 0.0125

B) Onset RT

Formula: Onset RT ~ Competition + (1 | Subject) + (1 | Quadruplet)

N= 2122

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	203.07	8.49	115.68	23.93	<.001
Competition Double	3.33	4.84	2010.78	0.68	.493

*alpha = 0.0125

C) Target Duration

Formula: Target Duration ~ Competition + (Competition | Subject) + (1 | Quadruplet)

N= 2122

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	145.89	3.96	118.39	36.83	<.001
Competition Double	0.04	2.08	102.49	0.02	.984

*alpha = 0.0125

D) Pre-target Duration

Formula: Pre-target Duration ~ Competition + (Competition | Subject) + (1 | Quadruplet)

N= 2040

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	141.65	3.61	144.33	39.23	<.001
Competition Double	4.35	1.41	84.95	3.09	.003

*alpha = 0.0125

Appendix D.3. Fixed effects of syntax models in Experiment 1 (sentence-dictation task)

Model formulas show random effects between brackets, with '‖' denoting uncorrelated random slopes.

A) Accuracy

Formula: Accuracy ~ Zipf + Word Type * Category + (Category ‖ Subject) + (Word Type | Quadruplet)

N= 9116

	Estimate	***SE***	***Z***	***p****
(Intercept)	1.86	0.09	21.33	<.001
Zipf	0.43	0.04	10.21	<.001
Word Type 1	0.29	0.06	4.69	<.001
Category 1	0.00	0.08	-0.06	.956
Word Type 1 : Category 1	-0.02	0.06	-0.25	.802

*alpha = 0.0125

B) Onset RT

Formula: Onset RT ~ Zipf + Word Type * Category + (Zipf + Word Type + Category ‖ Subject) + (Word Type | Quadruplet)

N= 4667

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	202.85	7.97	109.90	25.46	<.001
Zipf	4.50	2.19	206.74	2.05	.042
Word Type 1	-1.56	2.31	42.46	-0.68	.503
Category 1	1.98	5.77	40.72	0.34	.733
Word Type 1 : Category 1	0.44	2.12	35.89	0.21	.837

*alpha = 0.0125

C) Target Duration

Formula: Target Duration ~ Zipf + Word Type * Category+ (Zipf | Subject) + (Word Type | Quadruplet)

N= 4716

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	145.86	3.52	145.48	41.40	<.001
Zipf	-5.85	0.89	201.67	-6.54	<.001
Word Type 1	-0.71	1.16	36.95	-0.62	.542
Category 1	2.05	1.79	38.92	1.14	.259
Word Type 1 : Category 1	-1.07	1.15	36.80	-0.93	.360

*alpha = 0.0125

D) Pre-target Duration

Formula: Pre-target Duration ~ Zipf + Word Type * Category + (Zipf + Category || Subject) + (1 | Quadruplet)

N= 4484

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	142.14	3.54	151.55	40.11	<.001
Zipf	-0.42	0.48	207.53	-0.87	.387
Word Type 1	-1.33	0.36	4290.68	-3.71	<.001
Category 1	-0.35	1.78	39.35	-0.20	.844
Word Type 1 : Category 1	0.12	0.36	4306.66	0.35	.730

*alpha = 0.0125

Appendix E.1. Fixed effects of H vs. C models in Experiment 2 (question-answering task)

Model formulas show random effects between brackets, with '‖' denoting uncorrelated random slopes.

A) Accuracy

Formula: Accuracy ~ Zipf + Word Type + (Zipf | Subject) + (Word Type | Quadruplet)

N= 8746

	Estimate	*SE*	*Z*	*p**
(Intercept)	1.85	0.10	19.21	<.001
Zipf	0.19	0.05	4.30	<.001
Word Type H	-0.36	0.08	-4.33	<.001

*alpha = 0.0125

B) Onset RT

Formula: Onset RT ~ Zipf + Word Type + (Word Type | Subject) + (Word Type | Quadruplet)

N= 2858

	Estimate	*SE*	*df*	*t*	*p**
(Intercept)	234.26	10.08	113.22	23.24	<.001
Zipf	-13.28	4.10	635.22	-3.24	.001
Word Type H	-0.15	7.86	32.95	-0.02	.985

*alpha = 0.0125

C) Target Duration

Formula: Target Duration ~ Zipf + Word Type + (Word Type | Subject) + (Word Type | Quadruplet)

N= 2816

	Estimate	*SE*	*df*	*t*	*p**
(Intercept)	152.48	4.48	103.29	34.03	<.001
Zipf	-3.79	1.51	1596.18	-2.51	.012
Word Type H	0.31	3.74	37.93	0.08	.934

*alpha = 0.0125

D) Pre-target Duration

Formula: Pre-target Duration ~ Zipf + Word Type + (1 | Subject) + (Word Type | Quadruplet)

N= 2867

	Estimate	*SE*	*df*	*t*	*p**
(Intercept)	158.82	4.19	149.44	37.93	<.001
Zipf-score	-2.45	1.07	1193.97	-2.30	.022
Word Type H	2.36	1.49	44.17	1.58	.121

*alpha = 0.0125

Appendix E.2. Fixed effects of lexical competition models in Experiment 2 (question-answering task)

Model formulas show random effects between brackets, with ‘‖’ denoting uncorrelated random slopes.

A) Accuracy

Formula: Accuracy ~ Competition + (1 | Subject) + (1 | Quadruplet)

N= 4225

	Estimate	***SE***	***Z***	***p****
(Intercept)	1.39	0.10	13.53	<.001
Competition Double	0.16	0.08	2.05	.041

*alpha = 0.0125

B) Onset RT

Formula: Onset RT ~ Competition + (1 | Subject) + (1 | Quadruplet)

N= 1336

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	222.04	10.81	116.66	20.54	<.001
Competition Double	18.69	8.50	1256.24	2.20	.028

*alpha = 0.0125

C) Target Duration

Formula: Target Duration ~ Competition + (1 | Subject) + (1 | Quadruplet)

N= 1313

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	147.38	4.67	115.74	31.56	<.001
Competition Double	10.40	2.99	1215.52	3.48	<.001

*alpha = 0.0125

D) Pre-target Duration

Formula: Pre-target Duration ~ Competition + (Competition | Subject) + (1 | Quadruplet)

N= 1336

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	157.23	3.96	131.07	39.68	<.001
Competition Double	4.97	2.26	982.03	2.20	.028

*alpha = 0.0125

Appendix E.3. Fixed effects of syntax models in Experiment 2 (question-answering task)

Model formulas show random effects between brackets, with '‖' denoting uncorrelated random slopes.

A) Accuracy

Formula: Accuracy ~ Zipf + Word Type * Category + (Zipf + Category ‖ Subject) + (Word Type | Quadruplet)

N= 8746

	Estimate	***SE***	***Z***	***p****
(Intercept)	1.68	0.08	20.46	<.001
Zipf	0.21	0.04	4.66	<.001
Word Type 1	0.18	0.04	4.30	<.001
Category 1	0.11	0.07	1.58	.113
Word Type 1 : Category 1	-0.02	0.04	-0.59	.555

*alpha = 0.0125

B) Onset RT

Formula: Onset RT ~ Zipf + Word Type * Category + (Word Type | Subject) + (Word Type | Quadruplet)

N= 2858

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	234.20	9.31	125.05	25.16	<.001
Zipf	-12.92	4.11	634.06	-3.14	.002
Word Type 1	0.04	3.92	32.80	0.01	.991
Category 1	-4.66	4.95	35.69	-0.94	.353
Word Type 1 : Category 1	1.80	3.85	33.94	0.47	.644

*alpha = 0.0125

C) Target Duration

Formula: Target Duration ~ Zipf + Word Type * Category + (Word Type | Subject) + (Word Type | Quadruplet)

N= 2816

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	152.66	4.00	125.59	38.21	<.001
Zipf	-4.18	1.50	1409.81	-2.79	.005
Word Type 1	-0.15	1.84	38.06	-0.08	.935
Category 1	6.02	2.39	32.96	2.52	.017
Word Type 1 : Category 1	-1.41	1.81	37.01	-0.78	.442

*alpha = 0.0125

D) Pre-target Duration

Formula: Pre-target Duration ~ Zipf + Word Type * Category + (Category || Subject) + (Word Type | Quadruplet)

N = 2867

	Estimate	***SE***	***df***	***t***	***p****
(Intercept)	160.03	4.08	149.53	39.18	<.001
Zipf	-2.31	1.07	1253.19	-2.16	.031
Word Type 1	-1.17	0.74	43.78	-1.58	.121
Category 1	-2.79	2.07	39.83	-1.35	.184
Word Type 1 : Category 1	0.52	0.74	44.15	0.70	.491

*alpha = 0.0125