TOWARDS A SEMANTICS FOR LETTER PREDICATES*

Mathieu Paillé
University of Calgary and Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS)

1. Introduction

Work on predicate meanings often studies a particular class of predicates; for example, there is a rich and well-established literature on gradable predicates, and some work focuses on the distinction between natural-kind and artifact predicates (e.g. Hampton et al. 2009, Sassoon and Fadlon 2017). This paper aims to fill a gap in the literature by discussing predicates denoting symbols, like letters or written numbers. These have received almost no attention; I only know of Gasparri 2019 on written numbers.[1]

I will focus on letter predicates: words like *'f'* /ɛf/. As a first reason to find them interesting, consider that they apparently can be true of either graphemes or phonemes:

(1) a. (Phonologically,) there are two 'f's in 'philosophy.'
    b. (Orthographically,) there are two 'f's in 'traffic.'

This is a rather sharp case of polysemy. Letter terms can also be used as arguments (2), but I will focus on their predicative use.

(2) 'B' is the second letter of the alphabet.

I will show that the different interpretations of the letter predicate in (1) are not due to lexical ambiguity; rather, they arise from a single underlying lexical meaning.

This paper provides letter predicates with a lexical meaning that is context-sensitive in two ways. First, I integrate cross-linguistic variation in letters' phonemic values by having letter predicates take a particular writing convention as a contextual parameter. Then I integrate the phoneme–grapheme ambiguity via another parameter. This will lead to a lexical meaning for letter predicates that is more complex than what one might have imagined. But interestingly, I will then show that this meaning does not capture all the data. In particular, sentences I will call 'letter–letter co-predications' like (3) will not follow from

*I thank Alan Bale, Nina Haslinger, David Liebesman, Bernhard Schwarz, Benjamin Spector, and the audiences at the 2024 Alberta Conference on Linguistics (ACOL) at the University of Calgary and CLA 2025 at McGill.

[1]Symbol predicates are arguably a subclass of artifact predicates, but they are a subclass that has not been integrated in previous work on artifact predicates. Semiology, the study of signs (which include symbols), is an established field (e.g. Chandler 2007), but what I aim to do in this paper is to study the linguistic meanings of words denoting symbols, not symbols themselves.

Actes du congrès annuel de l'Association canadienne de linguistique 2025.
Proceedings of the 2025 annual conference of the Canadian Linguistic Association.

my proposal:

(3) In this language, the ‘d’s are ‘g’s.
$\rightsquigarrow$ ⟨d⟩ is pronounced as /g/ or /dʒ/; or
$\rightsquigarrow$ /d/ is written as ⟨g⟩

As such, the paper ends with an invitation for future work.

This paper is organized as follows. Section 2 focuses on (1) and suggests a first hypothesis about letter predicates' meanings. Section 3 then modifies the proposed meaning to integrate the two layers of context-sensitivity just described. Section 4 discusses co-predications of letter predicates like (3); it sketches out how one might try to make the hypothesis from section 3 predict such co-predications, but then shows that this would come with undesired consequences. Finally, section 5 turns briefly to another set of symbol predicates, namely numerals; it builds on work by Gasparri (2019) to give some context-sensitivity to their lexical entries.

2. An initial hypothesis for letter predicates

This section is an initial foray into letter predicates, focusing on data like (1) or (4).

(4) There is one ‘f’ in this word.
  a. $\rightsquigarrow$ one ⟨f⟩
  b. $\rightsquigarrow$ one /f/

I first discuss uses of letter predicates qua predicates of phonemes, then qua predicates of graphemes.

But first, a brief comment on the data in this paper. While linguists and non-linguists may have different intuitions about sentences like (1) or (4), I will abstract away from this. For example, sentences like (1a), repeated in (5a), might be obscure to non-linguists. They might instead use a sentence like (5b).

(5) a. (Phonologically,) there are two ‘f’s in ‘philosophy.’
  b. As far as sound is concerned, the word *philosophy* has two ‘f’s in it.

It is normal for experts and non-experts to talk differently about a topic. But crucially, the sentences in (5) are equivalent for our purposes, since in either case *‘f’* refers to a phoneme rather than a grapheme. Hence, I will put aside this kind of distinction.

2.1 Letters as phonemes

We have seen that letter predicates can sometimes be used to refer to a phoneme (6a). In fact, letter predicates can be applied to individuals explicitly stated to be sounds (6b).

(6) a. There are two ‘f’s in ‘philosophy.’
b. That sound is an ‘f.’

I assume sounds cannot have visual forms, so (6b) cannot mean that the sound is an occurrence of the grapheme ⟨f⟩.

Focusing on this, (7) starts working toward a lexical meaning for letter predicates.

(7) $[\![\text{‘f’}]\!] = \lambda x.\textbf{occurrence-of}(x, /\text{f}/)$.

(7) uses the meta-language relation **occurrence-of**. Occurrences are neither types nor tokens (Wetzel 2018). A given phonological ‘f’ is not the type /f/, nor is it necessarily just a token of /f/—tokens are spatiotemporally concrete entities. Many ‘f’s are not spatiotemporally concrete, as in (6a) (a statement about the word *philosophy*, an abstract entity). So we say there are two ‘occurrences’ of /f/ in *philosophy*.

(7) immediately raises the question of how to define more phonologically complicated letters. After all, one can say things like (8) about letters like ‘a’ that are associated with multiple phonemes:

(8) Phonologically, there are two ‘a’s in *dance bar*.

The apparently obvious (but not particularly insightful) way to deal with this is simply to posit disjunctive meanings:

(9) $[\![\text{‘a’}]\!] = \lambda x.$
$\textbf{occurrence-of}(x, /\text{æ}/) \vee \textbf{occurrence-of}(x, /\text{ɑ}/) \vee \textbf{occurrence-of}(x, /\text{eɪ}/)$.

The assumption here is that English speakers’ lexical meanings for letter predicates is based on English phonemes. I return to this assumption in section 3. There is also an open question of how to deal with speech sounds like schwa: schwa is often the pronunciation of ⟨a⟩ but not all schwas are describable as *an ‘a,’* since schwa can be spelled with other letters too. I also return to this in section 3.

2.2 Letters as graphemes

We have already seen data motivating that (7)/(9) constrain the meaning of letter predicates too much, by forcing the PHONEME meaning onto these predicates. Consider again (1b), repeated here:

(10) There are two ‘f’s in ‘traffic.’
a. $\approx$ two occurrences of ⟨f⟩ (true meaning)
b. $\not\approx$ two occurrences of /f/ (possible but false meaning)

(11) gives more examples of letter predicates denoting graphemes rather than phonemes:

(11) a. There are two ‘p’s in ‘philosophy.’
b. There is a ‘q’ is ‘qi.’
c. There is an ‘l’ in ‘salmon.’

All these sentences are true despite there being no /p/, /k(w)/ or /l/ in these words, respectively. How can we integrate this flexibility into the lexical meaning of letter predicates?

As a first attempt, we could posit lexical ambiguity between the PHONEME-meaning and the GRAPHEME-meaning of letter predicates:

(12) $[\![\text{‘f’}]\!] =$
a. $\lambda x.\mathbf{occurrence\text{-}of}(x, \langle \text{f} \rangle)$. ‘GRAPHEME meaning’
b. $\lambda x.\mathbf{occurrence\text{-}of}(x, /\text{f}/)$. ‘PHONEME meaning’

However, this cannot be right. It is sometimes possible to count occurrences of a letter by including both graphemes and phonemes:

(13) a. SCENARIO: *As part of a modern art exhibit, an artist sets up a room where there is nothing but a canvas with a big ‘f’ painted on it, and a speaker continuously playing a recorded* /f/.
b. I dislike both ‘f’s in this room.

If there was lexical ambiguity as postulated in (12), (13b) would be infelicitous in that scenario: ‘f’ could only be interpreted on the GRAPHEME meaning or the PHONEME meaning. Either way, there would only be one ‘f.’

Instead of ambiguity, we could instead posit a disjunctive meaning for letter terms:

(14) $[\![\text{‘f’}]\!] = \lambda x.\mathbf{occurrence\text{-}of}(x, \langle \text{f} \rangle) \vee \mathbf{occurrence\text{-}of}(x, /\text{f}/)$.

An issue at the moment is how to constrain (14) so that in a particular sentence, a speaker can use a letter predicate to refer *exclusively* to graphemes or phonemes. We return to this in section 3.2.

3. Two layers of context-sensitivity

This section builds on (14) by maintaining a disjunctive meaning for letter predicates, while adding two contextual parameters to their meaning. First, I will have letter predicates take a writing-relation argument. This is what will provide the phonemic values associated with a particular grapheme, rather than the phoneme being specified lexically. The grapheme denoted by a letter predicate will still be specified lexically. Second, the context-sensitivity between the GRAPHEME and PHONEME meanings will be obtained through a parameter constraining the extension of letter predicates to the writing relation's phonemes, graphemes, or both.

3.1 The first type of context-sensitivity: writing-system relations

Lexicalizing reference to phonemes as in (14) is a questionable move. One way to see this comes from the way letters' associated phonemes vary from one language to another.

Let me first set the stage with (15). This example is not problematic for (14), since it can be captured through the GRAPHEME disjunct.

(15) In French, 'r's are pronounced as /ʁ/.[2]
≈ 'In French, occurrences of the grapheme ⟨r⟩ are given the pronunciation /ʁ/.'

In (15), the letter predicate *'r'* denotes the grapheme ⟨r⟩. However, it is also possible to construct sentences where a letter predicate is used to denote not a grapheme but a phoneme of another language. Such sentences *are* problematic for (14). (16) gives some example sentences again with 'r' in French; unlike (15), in these sentences *'r'* is given the PHONEME-meaning. The examples rely on the fact that the Arabic letter ghayn ⟨غ⟩ is pronounced /ʁ/, like ⟨r⟩ in French.

(16) a. Arabic speakers have a French 'r' in their native language.
b. Ghayn (in Arabic) is a French 'r.'

To see why (16) is problematic, consider that so far, in modelling English speakers' meanings for letter terms, I have been assuming reference to English phonemes. On this view, *'r'* means (17) (cf. (14)):

(17) $[\![\text{`r'}]\!] = \lambda x.\textbf{occurrence-of}(x, \langle \text{r} \rangle) \vee \textbf{occurrence-of}(x, /\text{ɹ}/)$.

Yet, the meaning of *(French) 'r'* in (16) can be captured through neither disjunct in (17). The sentences in (16) state that Arabic speakers have a /ʁ/ in their phonology, not that they have /ɹ/ phonologically or ⟨r⟩ orthographically (they have neither of these). What we need is reference to the phoneme /ʁ/.

It is true that *'r'* in (16) is modified by *French*. But modification by *French* does not help matters; in fact, with the lexical meaning in (17), it is not clear how *'r'* and *French* could compose in a way that yields the observed meaning. If anything, one would expect *French 'r'* to necessarily pick out the GRAPHEME-meaning, since French writing contains occurrences of ⟨r⟩ but French phonology does not contain occurrences of /ɹ/. The sentences in (16) would then both be false, since Arabic orthography does not have ⟨r⟩.

I assume that we should not lexicalize reference to every phoneme a letter can be

[2]The pronunciation I have in mind for this sentence is roughly /ɑɹz ɑɹ pɹənaʊnst æz ʁː/.

associated with across all languages; speakers do not revise their lexical entries for letter predicates when they learn of a language associating the grapheme with a different phoneme. Thus, rather than attempting to lexicalize the cross-linguistic variation in pairings of graphemes and phonemes, I will take knowledge of grapheme–phoneme pairing to be fed semantically into letter predicates in the form of a contextually provided relation.

Assume speakers are familiar with one or more WRITING CONVENTIONS—anything from e.g. Italian writing conventions (18) to more specific things like the convention for Mandarin borrowings in English, where ⟨q⟩ produces /tʃ/.[3] The domain D_W of the writing-convention relation W is the set of graphemes in that writing system; the range R_W is the set of phonemes.

(18) **Italian writing convention**

$$\begin{array}{lcl} \langle \text{a} \rangle & \longrightarrow & /\text{a}/ \\ \langle \text{b} \rangle & \longrightarrow & /\text{b}/ \\ \langle \text{g} \rangle & \longrightarrow & /\text{ɡ}/ \\ & \searrow & /\text{dʒ}/ \\ \dots & & \end{array}$$

Lexically, letters take a writing convention as their first argument, as in (19).

(19) $[\![\text{‘f’}]\!] = \lambda W : \textbf{writing-convention}(W).\lambda x.$
$\textbf{occurrence-of}(x, \langle \text{f} \rangle) \vee \exists y\, [\langle \text{f} \rangle W y \wedge \textbf{occurrence-of}(x, y)]\,.$

The first disjunct in (19) is unchanged from our original (14), repeated in (20). But in the second disjunct, (20) lexicalizes reference to particular phonemes, while (19) lets a writing system argument provide the phonemic value.

(20) $[\![\text{‘f’}]\!] = \lambda x.\textbf{occurrence-of}(x, \langle \text{f} \rangle) \vee \textbf{occurrence-of}(x, /\text{f}/).$

Note that the meaning in (19) is asymmetric: it centres the graphemic component of letter predicates, with the phoneme parasitic on that grapheme. Thus, ‘f’ on the GRAPHEME-meaning can only denote occurrences of ⟨f⟩, but on the PHONEME-meaning, ‘f’ can denote any phoneme represented by ⟨f⟩ in whatever writing system the speaker has in mind.

I conclude this subsection by returning to a point raised in section 2.1: some letters like ‘a’ are associated with more than one phoneme, including among others /æ/, /ɑ/, and /eɪ/ in English. A result of this is that it is not always clear how many ‘a’ sounds there are in a word. Consider (21) (the two examples phrase the same question in different ways),

[3] Writing conventions are presumably more complicated than shown here; they should include information about digraphs (⟨ph⟩ as /f/), context-sensitive spelling rules, exceptions, etc.

with ‘apricot’ being pronounced /æpɹəkɑt/.[4]

(21) a. How many ‘a’ sounds are there in ‘apricot’?
b. Phonologically, how many ‘a’s are there in ‘apricot’?

Very reasonable answers to (21) include ‘one’ (due to /æ/ being written with an ⟨a⟩) and ‘two’ (/æ/ and /ɑ/), but the answers ‘zero’ and ‘three’ are not so bad as to be false. One could say zero since there is no /eɪ/, on the pronunciation I gave; and three due to the presence of /æ/, /ɑ/, and /ə/. To appreciate this possibility, imagine someone answering with *technically*:

(22) Technically, {zero, three}.

Technically here signals that the speaker is referring to a particular writing convention that they consider counter-intuitive or unusual, perhaps due to appealing to specialist knowledge in including schwa, or due to using an overly conservative writing convention that only ties ⟨a⟩ to /eɪ/. Either way, the four different answers to (21) correspond to different values for the writing system parameter.

Let me elaborate on this claim that different writing system parameters are available in the context of discussing an English word. Given the mess of English orthography, it is not outrageous to think that speakers can make up ad hoc ways of thinking of English writing conventions. The phonemic values of some letters could be associated with those letters as a rule, or as an exception in a particular word. To cite an extreme example, the fact that the second ⟨f⟩ in ‘fifth’ is silent for some speakers presumably is not due to the possibility that $(\langle \text{f} \rangle, \varnothing) \in W_e$ (W_e being an English writing convention), but rather because $(\langle \text{fifth} \rangle, /\text{fɪθ}/) \in W_e$. Likewise, one can take ⟨a⟩ to produce /æ/ (or /eɪ/, or /ɑ/) conventionally and have other values due to word-specific information in W_e; or one can take ⟨a⟩ to represent more than one phoneme conventionally. Thus, for example, the answer ‘one’ to the question in (21) takes /æ/ to be the conventionalized phoneme for ⟨a⟩ in the ad hoc value for W in this discourse context. In this case, the speaker’s choice for W_e is presumably influenced by the context at hand, in which only /æ/ is represented by an ⟨a⟩ grapheme.

3.2 The second type of context-sensitivity: Phoneme-only, grapheme-only uses of letter predicates

(19) has added context-sensitivity in letter predicates’ meanings in letting them be used vis-à-vis a particular writing convention. But as we have seen, there is another kind of context-sensitivity, namely that letter predicates can be used exclusively for phonemes (23a), for graphemes (23b), or for both (23c):

[4] I treat the second vowel as a phonemic schwa for simplicity.

(23) a. There is one ‘f’ in ‘traffic.’
b. There are two ‘f’s in ‘traffic.’
c. I dislike both ‘f’s in this room. *(in the art-installation scenario, (13a))*

While (19) captures (23c), it cannot capture (23a–b), where the speaker is only counting one kind of occurrence of ‘f,’ namely occurrences of ⟨f⟩ or of /f/. For this reason, (24) gives our lexical entry a second piece of context-sensitivity. The variable S is a set corresponding to the domain of the writing convention, to its range, or to the union of these two. It lets speakers limit the extension of ‘f’ to graphemes, phonemes, or both.[5]

(24) $[\![\text{‘f’}]\!] = \lambda W : \textbf{writing-convention}(W).\lambda S : S = D_W \vee S = R_W \vee S = D_W \cup R_W.\lambda x.$
$[\langle \text{f} \rangle \in S \wedge \textbf{occurrence-of}(x, \langle \text{f} \rangle)] \vee \exists y\, [y \in S \wedge \langle \text{f} \rangle W y \wedge \textbf{occurrence-of}(x, y)]\,.$

For example, by choosing to identify S with R_W, the speaker limits the denotation of ‘f’ to occurrences of /f/ (or whatever phoneme ⟨f⟩ represents in the relevant writing system).

(24) is the final proposal in this paper.

4. A remaining puzzle: letter–letter co-predications

Even with the somewhat complicated semantics in (24), we cannot obtain all the data. This section shows why, but without providing a new hypothesis.

The problem is due to a class of data I will call ‘letter–letter co-predications’: cases where two letter predicates are predicated of the same individual. We already saw one such example in (16b):

(25) Ghayn (in Arabic) is a French ‘r.’

The meaning of (25) is that the grapheme ⟨غ⟩, called ghayn, represents /ʁ/—call this the PRONOUNCED-AS meaning of letter–letter co-predications.

Stepping back from this particular example, we observe in letter–letter co-predications not just the PRONOUNCED-AS meaning of (25), but also a WRITTEN-AS meaning:

[5] A problem at the moment, which I will not fix in this paper, is how to rule out (i):

(i) #There are three ‘f’s in ‘traffic.’

After all, there are two orthographic ‘f’s and one phonological ‘f.’ See Liebesman and Magidor 2025 for extensive discussion of similar problems. Informally, the difference between (i) and the art-installation scenario (13a)/(23c) is that there is a causal connection between the graphemic and phonemic ‘f’s in (i), but not in (13a); counting must involve distinct elements.

(26) In this language, the ‘d’s are ‘g’s.

a. $\rightsquigarrow$ $\langle$d$\rangle$ is pronounced as /g/
b. $\rightsquigarrow$ /d/ is written as $\langle$g$\rangle$

To reiterate from section 2: I assume that an occurrence of a grapheme cannot ‘be’ an occurrence of a phoneme, or vice-versa; they are different kinds of entities.

To appreciate the difficulty of these data, I start by sketching out a first hypothesis in section 4.1; it is conservative in preserving the meaning for letter predicates given in (24). The hypothesis is not fully worked out, but it attempts to identify a path forward. I then show in section 4.2 that it makes a fundamentally wrong prediction for definite expressions (*the ‘d’s* in (26)).

4.1 A hypothesis based on PHONEME meanings

The hypothesis I sketch out in this section attempts to capture the truth conditions of examples like (26) by relying entirely on the PHONEME meaning of letter predicates, for both the PRONOUNCED-AS meaning and the WRITTEN-AS meaning. This is arguably counter-intuitive: the way one would paraphrase the meanings of (26) involves referencing one grapheme and one phoneme, not two phonemes. In section 4.2, I will show that this counter-intuitiveness is well-grounded; the current hypothesis relying on two PHONEME meanings appears empirically incorrect.

Let us start with the PRONOUNCED-AS meaning:

(27) In this language, the ‘d’s are ‘g’s.
$\rightsquigarrow$ $\langle$d$\rangle$ is pronounced /g/

Call the language referred to in (27) Language L. Let W_e be the writing convention for English and W_l be the writing convention for Language L. For *‘g’s*, we feed *‘g’* the English writing convention and a set $S = R_{W_e}$ consisting of English phonemes.

(28) $[\![[\text{‘g’ } W_e]\ R_{W_e}]\!] = \lambda x.\exists y[y \in R_{W_e} \wedge \langle \text{g} \rangle W_e y \wedge \textbf{occurrence-of}(x,y)]$.
($\approx x$ is an occurrence of a phoneme associated with $\langle$g$\rangle$ in English)[6]

We also use the same PHONEME meaning for *‘d’* in *the ‘d’s*, by feeding it the set R_{W_l}:

(29) $[\![[\text{‘d’ } W_l]\ R_{W_l}]\!] = \lambda x.\exists y[y \in R_{W_l} \wedge \langle \text{d} \rangle W_l y \wedge \textbf{occurrence-of}(x,y)]$.

Now we must let (28) and (29) compose. As stated above, this hypothesis is not fully worked out, but the idea would be to use an equative copula to arrive at a meaning like: “the

[6] I have suppressed from (28) the left disjunct—the GRAPHEME meaning—for simplicity of presentation. That disjunct is irrelevant once *‘g’* has been fed the set corresponding to English phonemes, preventing it from referring to a grapheme.

phoneme represented by ⟨d⟩ in Language L = the phoneme represented by ⟨g⟩ in English." Thus, on this hypothesis, even if it feels like (27) is a statement about the grapheme ⟨d⟩ (asserting it represents /g/), the truth conditions arise from the PHONEME meaning of both letter predicates.

The same goes for the WRITTEN-AS meaning, repeated here:

(30) In this language, the 'd's are 'g's.
$\rightsquigarrow$ /d/ is written ⟨g⟩

(30) equates the phoneme of English ⟨d⟩ with the phoneme of the ⟨g⟩ of Language L—hence meaning that ⟨g⟩ is pronounced /d/. More formally, this means that the only difference from the PRONOUNCED-AS meaning is that we switch which letter predicate is fed W_e/R_{W_e} and which one is fed W_l/R_{W_l}:

(31) a. $[\![[\text{'d'}\ W_e]\ R_{W_e}]\!] = \lambda x.\exists y[y \in R_{W_e} \wedge \langle \text{d} \rangle W_e y \wedge \textbf{occurrence-of}(x,y)]$.
b. $[\![[\text{'g'}\ W_l]\ R_{W_l}]\!] = \lambda x.\exists y[y \in R_{W_l} \wedge \langle \text{g} \rangle W_l y \wedge \textbf{occurrence-of}(x,y)]$.

On this hypothesis, then, it is the PHONEME meanings of letter predicates that are equated, in both the WRITTEN-AS and the PRONOUNCED-AS meanings.

4.2 A problem: definites show the presence of GRAPHEME meanings

I repeat that the above hypothesis is not fully worked out. But simply based on what has been stated so far, it is already possible to see that the hypothesis gives rise to at least one empirical problem. Let us first simplify things a bit by considering a sentence with singular predication, and focus just on the PRONOUNCED-AS meaning (nothing hinges on this):

(32) The 'd' is a 'g.'
$\rightsquigarrow$ ⟨d⟩ is pronounced /g/

Now notice that, working with the above meanings, in particular (29), we obtain (33) for *the 'd'*:

(33) $[\![\text{the}\ [[\text{'d'}\ W_l]\ R_{W_l}]]\!] = \iota x.\exists y[y \in R_{W_l} \wedge \langle \text{d} \rangle W_l y \wedge \textbf{occurrence-of}(x,y)]$.

This denotes the unique individual that is an occurrence of the phoneme represented by ⟨d⟩ in Language L—so an occurrence of a /g/.

What we expect, then, is for the definite *the 'd'* on the PHONEME meaning in (33) to identify the unique occurrence of /g/ in a particular context. However, this is not in fact correct: *the 'd'* can be used even if there is more than one phonemic /g/. Imagine that in Language L, for historical reasons, it is not only ⟨d⟩ but also ⟨g⟩ that is used to represent /g/. Now consider a word spelled ⟨gid⟩, pronounced /gig/:

(34) a. SCENARIO: *You are discussing a word in Language L spelled* ⟨gid⟩ *and pronounced* /gig/.
b. The ‘d’ is a ‘g.’

In this scenario, it is true that there is a unique ⟨d⟩ grapheme, but it is not true that there is a unique /g/ phoneme. Yet, (32)/(34b) is felicitous in that scenario. (33) therefore cannot be the right meaning for *the ‘d’*.

Hence, the above hypothesis for letter–letter co-predications is presumably on the wrong track. I wrote in section 4.1 that the hypothesis was counter-intuitive in building on the PHONEME meaning of both letter predicates in (32), when intuitively the sentence makes a statement about the grapheme ⟨d⟩. We have now seen that this is truly a problem for the hypothesis. In light of this, letter–letter co-predications remain an open puzzle for letter predicate meanings.

5. Context-sensitivity in number-symbol expressions

The above concludes my discussion of letter predicates. In this last section, I extend some of my claims to another set of symbol expressions, namely number terms.

I mentioned in section 1 that there is prior work on symbol predicates by Gasparri (2019), who discusses numbers. Number terms can be used for numerals or for symbols:

(35) a. The number fifteen is odd. (Gasparri 2019: 564, 568)
b. The number fifteen has two digits.

Being odd is a property of numbers, not symbols; and *having two digits* is a property of symbols, not numbers.

The semantics Gasparri gives for number terms is:

(36) $[\![\text{the number fifteen}]\!] = 15 \bullet \langle 15 \rangle$.

In (36), ‘15’ refers to the number fifteen while ‘⟨15⟩’ refers to the digraphic symbol. As for the dot, this is the ‘dot object’ approach (e.g. Asher 2011) to category-mismatch co-predications, sentences where two predicates that one would think hold of categorically different types of entities are both applied to the same entity.[7] (37) gives an example of a category-mismatch co-predication with number terms:

(37) The number fifteen is odd and has two digits. (Gasparri 2019: 565)

The right approach to category-mismatch co-predications is still being debated (see e.g. Liebesman and Magidor 2025 and citations therein for recent work on this); I will simply

[7]Category-mismatch co-predications are usually simply called ‘co-predications’ in the literature.

follow Gasparri's dot notation for the purposes of this section.[8]

What matters about (36) for our purposes is that reference to $\langle 15\rangle$ is part of the phrase's meaning. Gasparri (2019: 568) notes that this is contextually variable, since there are many ways to write a particular number. Fifteen in binary is $\langle 1111\rangle$, for example. Hence, among binary enthusiasts, the meaning of *the number fifteen* would be:

(38) $[\![\text{the number fifteen}]\!] = 15 \bullet \langle 1111\rangle$.

Gasparri (2019) is not explicit on how (38) comes about, so I suggest to formalize this as a kind of context-sensitivity parallel to what I formalized for letter terms. Much like letter terms' meanings arise in part by taking a writing-convention as an argument, number terminology must also have a contextually provided relation from numbers to written symbols. To avoid getting into the composition, let me just assign this variable to the entire phrase *the number fifteen* (39). Call the relation from numbers to symbols (containing ordered pairs like $(15,\langle 15\rangle)$ in decimal or $(15,\langle 1111\rangle)$ in binary) a 'numeral convention' N. Numeral conventions are specifically functions, assuming that every number is only written in one way in a given convention (unlike letters, which potentially have multiple phonemic associations).

(39) $[\![\text{the number fifteen}]\!] = \lambda N : \textbf{numeral-convention}(N).15 \bullet \iota x[15Nx]$.

On the right side of the dot, we have the unique symbol associated with a number according to the numeral convention.[9]

6. Conclusion

I have shown that letter predicates are context-dependent in at least two ways, taking both a writing convention and a variable corresponding to whether their denotation includes graphemes, phonemes, or both. This paper suggested to deal with this by putting two contextual parameters in letter predicates' lexical meanings.

[8]We can also find category-mismatch co-predications with letter terms:

(i) The 'a' in *Ann* is capitalized and nasal.

After all, *capitalized* is a predicate of graphemes while *nasal* is a predicate of phonemes. Needless to say, the proposal in (24) will need to be amended to cover these kinds of examples.

[9]What I do not find entirely obvious is on what basis sentences like (i) are ruled out:

(i) #The number 15 has seven letters.

After all, one way to write numbers is by spelling them out linguistically; one can write 15 as *fifteen*, 2 as *two*, and so on. To deal with this, we have to claim that for whatever reason, this is simply not a 'numeral convention.'

A number of interesting questions remain. Most obviously, I described in section 4 difficulties capturing letter–letter co-predications on the current proposal. Another remaining challenge is that of category-mismatch co-predications, mentioned in section 5 and in particular fn. 8. What seems certain is that the empirical domain of symbol predication is much more empirically interesting than one might naively expect.

References

Asher, Nicholas. 2011. *Lexical meaning in context: a web of words*. Cambridge: Cambridge University Press.

Chandler, Daniel. 2007. *Semiotics: the basics*. London: Routledge, 2nd ed.

Gasparri, Luca. 2019. A numeral oddity. *Journal of Semantics* 36: 563–571. https://doi.org/10.1093/jos/ffz006.

Hampton, James A., Gert Storms, Claire L. Simmons, and Daniel Heussen. 2009. Feature integration in natural language concepts. *Memory & Cognition* 37(8): 1150–1163.

Liebesman, David, and Ofra Magidor. 2025. *Property versatility and copredication*. Oxford: Oxford University Press.

Sassoon, Galit Weidman, and Julie Fadlon. 2017. The role of dimensions in classification under predicates predicts their status in degree constructions. *Glossa* 2(1): 1–40.

Wetzel, Linda. 2018. Types and Tokens. In *The Stanford encyclopedia of philosophy*, ed. Edward N. Zalta. Metaphysics Research Lab, Stanford University, Fall 2018 ed.