Approximation of Complex Numerals

Using Some

Curt Anderson
Michigan State University

1. Some Can Play a Role in Approximation

English *some* normally plays the role of a determiner, appearing before a noun phrase, as in (1). This use of *some* has most often caught the attention of linguists and philosophers. However, *some* can be used in a non-canonical way with numerals, as in (2) and (3).

(1) a. There were some dogs in the yard.
   b. Some man is crossing the street.
   c. I put some apple in the salad.
(2) a. Some twenty people attended the party.
   b. Some 5 million people are without health insurance.
(3) a. Twenty-some people were at the party.
   b. 5 million-some people are without health insurance.

The salient observation about the examples in (2) and (3) is that not only is *some* allowed to modify the cardinal number in a position before the number, but there exist cases where *some* can be in a modifier relationship with the number while appearing after it as well. Throughout the rest of the paper, I will call the former construction the pre-numeral *some* and the latter the post-numeral *some*.

The pre-numeral *some* is able to modify a variety of numerals, demonstrated in (4). However, quite mysteriously, the post-numeral *some* cannot modify some of these same numerals, as shown in (5).

(4) a. Some ten people attended the lecture.
   b. Some five years after an economic crisis
   c. The original text was written some twenty-five years ago.
(5) a. *Ten-some people attended the lecture.
   b. *Five-some students were arrested after the riot.
   c. *The original text was written twenty-five some years ago.
Finally, there are interpretational differences between these two non-canonical uses of *some* as well. In the pre-numeral *some*, the natural interpretation is one of approximation—values close to the number being modified by *some* are implicated in the meaning of the pre-numeral *some*. In contrast, with the post-numeral *some*, there is an “at least” interpretation—the values for the number implied in this construction start at the number being modified and continue up the scale.

A couple of questions naturally arise here. The first is how these two uses of *some* are related to each other, and whether they are the same *some*. Second, how does the semantic system build an approximative meaning for the pre-numeral *some* and an “at least” interpretation for the post-numeral *some*? Finally, what is the nature of the syntactic restrictions between the two *somes*? I explore and answer each of these questions, analyzing *some* as being sensitive to Hamblin alternatives (Hamblin, 1973). These alternatives are constructed in separate ways for the two *some* constructions at issue in this paper, with the pre-numeral *some* invoking imprecision alternatives, alternatives that model Lasersohnian pragmatic halos (Lasersohn, 1999; Morzycki, 2011), while the post-numeral *some* implies a covert *wh*-word that abstracts over positions in the syntax of cardinal numbers, providing numerical alternatives to *some*. Evidence for this covert *wh*-word comes from a similar construct to the post-numeral *some* in Japanese.

2. Post-numeral *Some* Is Sensitive to Numeral Syntax

To account for the post-numeral *some*, it’s useful to return to its interpretation and to its restrictions. What I will show here is that there is a common source for both of these, namely that syntactic structure in complex numbers explains both the syntactic restrictions of post-numeral *some* and its interpretation. The core idea will be that numbers are derived compositionally, and that the post-numeral *some* is sensitive to the structure of numerals.

The nature of the restrictions on post-numeral *some* strongly suggests that numbers have a complex syntactic structure. That numbers are built compositionally is not a new idea, having appeared at least as early as Hurford (1975), and more recently in Ionin and Matushansky (2006), Zweig (2005), and others. To start, we notice that not combinations of numerals are licit — not all numbers can appear in all syntactic positions. This is demonstrated in (6). In fact, in the absence of a word such as *eleven*, we might have otherwise predicted that *ten one* could have the same meaning as *eleven*, but in English it is simply ungrammatical.
(6) a. *Three five (intended: thirty five)
    b. *Ten one (intended: eleven)
    c. *Fifteen eight (intended: one hundred fifty eight)

However, numbers do combine with other numbers more generally. Twenty-five is composed of the two numbers twenty and five, for instance, while one hundred twenty five is composed of one hundred and twenty five. And this is of course recursive: twenty-five in one hundred twenty five is also built from twenty and five. The conclusion should be that complex numerals are built from smaller, less complex numerals.

What we notice about the post-numeral some is that it is sensitive to these same restrictions: the numbers in (6) cannot combine with the post-numeral some as well, shown in (7). The conclusion I draw is that the post-numeral some construction is sensitive to restrictions inherent in how complex numerals are constructed.

(7) a. *Three-some
    b. *Ten-some
    c. *Fifteen-some

Some additional evidence that the post-numeral some is sensitive to the syntactic structure of the numeral comes from decimal numbers. Decimal numbers in English, at least in casual speech, have a list-like structure to them, where they are simply a sequence of numbers (for instance, 1.634 is commonly uttered as one point six three four). The post-numeral some can abstract over parts of decimal numbers, provided there is a suitable context, as shown in (8).

(8) A student in a chemistry class need to fill a test tube with a quantity of fluid. The exact amount of fluid is 1.635 milliliters, but the student cannot remember this number. This student can say:
    I need to fill this with 1.63-some milliliters of fluid.

3. Approximation of Numerals in Japanese

Like English, Japanese builds larger, more complex numbers by putting together smaller numbers. As shown in (9a), Japanese juu-ichi “eleven” is built by putting together the morphemes juu “ten” and ichi “one.” Relatedly, in (9b), ni “two” and juu “ten” are put together to form the numeral ni-juu “twenty.”

(9) a. juu-ichi
    ten-one
    “eleven”

    b. ni-juu
    two-ten
    “twenty”
Like in the English post-numeral *some* construction, Japanese has a way of being imprecise about the precise value of some number. The example in (10) has an interpretation similar to the English post-numeral *some*: a syntactic position in the numeral has been abstracted over with *nan* “what” in order to build an “at least” interpretation. Unlike English, Japanese is more flexible in what may be abstracted over. This is demonstrated in (11), where, due to the Japanese equivalent of English “twenty” being composed of “two” and “ten,” the speaker can make assertions about some multiple of ten by using *nan* in the position that would otherwise be occupied by the numeral *ni* “two,” as in (9b).

(10) Juu-nan-nin-ka-ga kita
    ten-what-cl(people)-ka-nom came
    “10 plus x people came.”

(11) Nan-juu-nin-ka-ga kita
    what-ten-cl(people)-ka-nom came
    “x multiple 10 people came.”

What is interesting about (10) and (11) is how these approximative constructions are composed. In each, there is a morpheme that appears in the position of the number that is abstracted over, *nan*. *Nan* is an indeterminate pronoun, roughly equivalent to “what” in English. But, Japanese looks like English in these constructions in at least one other way, with the particle *ka* in (10) and (11); *ka* is sometimes analyzed as carrying existential force, similar to *some* (Slade, 2011; Cable, 2010; Kratzer & Shimoyama, 2002).

To tackle the English pre-numeral and post-numeral *somes* which are the focus of this paper, I suggest we should understand the Japanese constructions above first. Looking at the Japanese will help us construct an analysis of the English facts. Two theoretical pieces will be introduced here: the Hamblin semantics analysis of Japanese indeterminate pronouns of Shimoyama (2001) and Kratzer and Shimoyama (2002), and *ka* as denoting a choice-functional variable.

Kratzer and Shimoyama (2002) and Kratzer (2005) provide an analysis of Japanese indeterminate pronouns using a Hamblin alternative semantics (Hamblin, 1973). The idea behind an alternative semantics is that related sentential meanings can be represented in parallel to each other as sets of meanings. This is widely used for the semantics of questions, where questions denote alternatives representing answers to the question (Hamblin, 1973; Karttunen, 1977). Shimoyama (2001) suggests that indeterminate pronouns in Japanese, which resemble *wh*-words (question words, such as “who” and “what” in English), can be given an alternative semantics, where they directly denote sets of alternatives.

In an alternative semantics, a new notion of composition is needed, since sets themselves cannot be combined. The basic method of composition, Function
Application (Heim and Kratzer, 1998), is reformulated as Pointwise Function Application (see Kratzer and Shimoyama (2002) for details). The intuition is to apply each function in the first set of alternatives to each object in the second set of alternatives, yielding a new set. Throughout the course of the derivation, a set of alternatives will continue to expand, due to each successive application of Pointwise Function Application creating a larger set of alternatives from the alternatives for some function and some object.

Sets of alternatives must be captured and mapped to a single alternative, if a coherent declarative meaning is to be constructed. The intuition is that the particle *ka* associates with alternatives, stops alternatives from expanding, and maps the alternatives to a single alternative. Analyses such as Slade (2011) and Cable (2010) argue that *ka* denotes a choice functional variable, a variable for a function from sets to a member of a set. In effect, this is a way of providing existential quantification. At the level of the DP, as in examples (10) and (11), the role of *ka* would be to close off the set of alternatives and select a single alternative to project. This conception of *ka* has connections to the meaning of English *some* which, as an indefinite determiner, also seems to have existential force associated with it.

In Japanese, *ka* can also serve as a question particle. The question particle *ka* associates with these alternatives if there is no other intervening *ka* to capture the alternatives. If there is an intervening *ka* in the question, however, what we expect is for the question word to only be able to associate with the singleton alternative—that is, for there to be only a yes/no question interpretation. As shown in (12) and (13), this is what we find, where an intervening *ka*, as in (12), forces a yes/no question interpretation, but no intervening *ka* in (13) allows all the numerical alternatives to project. When the operator *ka* is present low in the structure, at the level of the DP, it stop the alternatives from the *wh*-word from expanding, forcing the yes/no question interpretation. When *ka* is not present at the DP level, the alternatives from the *wh*-word—the numerical alternatives associated with abstracting over part of the complex numeral—can continue to expand upward, until they are caught by the question particle *ka*. At that point, they are used in forming the question, a question that’s seeking information about which number of people came.

(12) Nan -juu -nin -ka -ga -kita ndesu ka?
   what -ten -cl(people) -ka -nom came be Q
   “Is it the case that x multiple 10 people came?” (yes/no question)

(13) Nan -juu -nin -ga kita ndesu ka?
   what -ten -cl(people) -nom -came be Q
   “What is the number x such that x multiple 10 people came?” (wh-question)

The Japanese data is important in a few key respects. First, it quite transparently
shows that numerals are constructed syntactically. Second, it suggests that there may be overt or covert morphemes that can be used to hook into the syntax of the numeral in order to build an approximate interpretation. Third, a connection between ka and some is made, as both contribute existential force. Finally, it suggests that there is a role to be played by alternatives in the computation of approximate numerical meanings.

4. Semantics for Approximation with Some

4.1. Post-numeral some

The analysis of the post-numeral some will consist of three components motivated in part by the Japanese data in the previous section. First, English numerals will also be syntactically complex. That English numerals are syntactically complex is argued for by Hurford (1975), Ionin and Matushansky (2006), and others. Second, like Japanese, I will assume that there exists a covert wh-word present in the post-numeral some that abstracts over a position in the numeral. Finally, I use an alternative semantics to model approximation. The wh-word will be interpreted in situ as a set of numerical alternatives, which will combine pointwise with the other alternatives composing the numeral.

As has been argued, large numerals have structure to them. A complex numeral such as sixty-five thousand two hundred forty five is constructed out of smaller numerals such as forty five, sixty-five thousand, and two hundred, each of which is also built of smaller numerals. Certain configurations of numerals are additive, while others are multiplicative. For instance, forty five is additive, as it is the number 45, which is simply the addition of 40 and 5, while sixty-five thousand is multiplicative, as it 65 multiplied by 1000. A complex number intuitively has a constituency like in (14), where configurations of numerals are combined via an additive or multiplicative process, as illustrated via a + or \times dominating the numerals being combined.

(14)
As argued previously, the post-numeral *some* is sensitive to the structure of the numeral. More specifically, it is only numerals that are composed additively that the post-numeral *some* can pick out for approximation. This is illustrated in (15) and (16), where *some* can be used in (15) since the meanings of the numbers are built from adding the two numbers together, while *some* cannot be used with the intended meaning in (16) since that meaning comes about from multiplying the numbers.

(15) a. Twenty-*some* people  
   b. Sixty-*some* thousand dollars

(16) a. *Two-*some people  
   Intended: twenty thousand people  
   b. *Sixty-*some dollars  
   Intended: sixty-thousand dollars

More generally, the meanings available when the post-numeral *some* can be used depend on what numbers can be licitly used additively with the modified number. In *twenty-*some, for instance, only numbers one through nine can be composed additively with *twenty*, and hence *twenty-*some has the interpretation of denoting a number between 21 and 29.

I set aside the precise internal structure of numerals for further work, as all that is crucial in this paper is that numerals have structure associated to them. I will represent numerals as simply XPs adjoined to NP, as in (17).

(17)

```
  NP  
  /   
XP   NP  
  / 
twenty   people
```

To model the semantics of the numeral, I assume a domain of numbers $D_n$. Addition $+$ and multiplication $\times$ are defined over pairs of numbers in $D_n$, with the result being another number in $D_n$. As I will be using an alternative semantics to model approximation, numerals themselves will denote sets of numbers rather than directly denoting numbers. Simple number words denote singleton sets whose member is a number in $D_n$. For instance, $[\text{twenty}]$ is just the set containing only the numeric value of twenty, $\{20\}$.

(18) $[\text{twenty}] = \{20\}$

(19) $[\text{five}] = \{5\}$
I assume a typeshift \text{CARD} to convert numerals to properties, type \langle e,t \rangle. This is defined in (20).

\begin{align}
(20) \; \llbracket \text{CARD} \alpha \rrbracket &= \lambda x \; [\; |x| = \llbracket \alpha \rrbracket \; ]
\end{align}

where \(\alpha\) is a numeral

In the post-numeral \textit{some} examples, based on the Japanese data, I suppose the existence of a covert \textit{wh}-word. This covert \textit{wh}-word essentially acts as a placeholder for any of the numerals that could have been composed with another numeral additively. In \textit{twenty-some}, the syntactic structure at LF (the level of representation responsible for semantic interpretation) would be \textit{twenty-WH-some}, where \text{WH} is the covert word and stands in for the numerals \(1\) through \(nine\). Hence, WH has the meaning of the set \(\{1, \ldots , 9\}\) in \textit{twenty-WH-some}. However, the meaning of the WH depends on the syntactic structure; \text{WH} could take on different values in other examples, such as denoting \(\{1, \ldots , 99\}\) in \textit{two hundred-WH-some}.

\begin{align}
(21) \; \text{twenty-WH-some} \; \llbracket \text{WH} \rrbracket &= \{1, \ldots , 9\} \\
(22) \; \text{two hundred-WH-some} \; \llbracket \text{WH} \rrbracket &= \{1, \ldots , 99\}
\end{align}

How do numbers get composed in this system? As we’re working with sets of numbers and not functions, the typical mode of semantic composition, Function Application (Heim & Kratzer, 1998), will not work. I propose Pointwise Addition and Pointwise Multiplication, derived from Kratzer and Shimoyama's Pointwise Function Application. These are defined in (23a) and (23b), respectively.

\begin{align}
(23) \; \text{a. Pointwise Addition} \\
\text{Where} \; \llbracket A \rrbracket \; \text{and} \; \llbracket B \rrbracket \; \text{are sets of numbers,} \\
\llbracket C \rrbracket &= \{c: \exists a \in \llbracket A \rrbracket \land \exists b \in \llbracket B \rrbracket \land c = a + b\} \\
\text{b. Pointwise Multiplication} \\
\text{Where} \; \llbracket A \rrbracket \; \text{and} \; \llbracket B \rrbracket \; \text{are sets of numbers,} \\
\llbracket C \rrbracket &= \{c: \exists a \in \llbracket A \rrbracket \land \exists b \in \llbracket B \rrbracket \land c = a \times b\}
\end{align}

The idea behind these rules is simple: everything from the first set is added or multiplied in turn with each item from the second set. With two singletons, this process is trivial; all that is to be done is to add (or multiply) the only item from the first with the only item from the second. With non-singleton sets, Pointwise Addition and Multiplication is much more interesting. In a case with non-singleton sets, each item from the first set will be added or multiplied with each item from the second set, generating a third set. This is what happens when a
number such as \( \{20\} \) is added pointwise to the set of numbers denoted by WH. This process is illustrated in (24).

(24) a. \( \llbracket \text{twenty WH} \rrbracket \)
    \[ = \llbracket \text{twenty} \rrbracket + \llbracket WH \rrbracket \]
    \[ = \{20\} + \{1, \ldots, 9\} \]
    \[ = \{20+1, 20+2, \ldots, 20+9\} \]
    \[ = \{21, 22, \ldots, 29\} \]

How does the numeral combine with the noun phrase in the cases without WH and *some*? For this, the CARD typeshift is involved to convert the number to a property. The property can then be combined intersectively with the denotation of the noun phrase.

(25)

(26) \( \llbracket \text{CARD twenty people} \rrbracket \)
    \[ = \{ \lambda x. \llbracket \text{CARD twenty} \rrbracket (x) \land \llbracket \text{people} \rrbracket (x) \} \]
    \[ = \{ \lambda x. |x| = 20 \land \text{people}(x) \} \]

The set-based representation pays off when we consider numerals with WH. The purpose of WH was to introduce a set of alternatives into the representation. When WH is in the numeral, the numeral will denote a non-singleton set of numbers, as shown above in (24). The role of *some* in the post-numeral *some* construction is to map this set of alternatives to a single alternative.

The method of doing this will be a choice function (Reinhart, 1997; Winter, 1997; Kratzer, 1998), a function from a set to a member of that set. The choice functional analysis for *some* can be developed as in (27), where the alternatives of the expression \( \alpha \), a placeholder for the numeral, are mapped to a singleton. The value of the choice functional variable \( f \) is supplied by the context.

(27) **Choice Functional Some (First Version)**

\[ \llbracket \text{some } \alpha \rrbracket = \{ f(\llbracket \alpha \rrbracket) \} \]

where \( f \) is a choice functional variable

The derivation of *twenty WH some people* would proceed as follows. WH combines with *twenty*, forming a set of numerical alternatives. *Some* selects
from among these alternatives, and the typeshift \textit{CARD} maps the number to a property. This property combines intersectively with the denotation of the NP, in the same fashion as in (26).

(28)

(29) a. $[[\text{twenty WH}]] = \{21, 22, \ldots , 29\}$

b. $[[\text{twenty WH some}]] = \{f([\text{twenty WH}])\}$

c. $[[\text{CARD} [\text{twenty WH some}]]] = \{\lambda x. |x| = f([\text{twenty WH}])\}$

d. $[[\text{CARD} [[\text{twenty WH some}]] \text{ people}]]$

To summarize, \textit{some} is sensitive to alternatives, picking from among alternatives by way of a choice function. A covert \textit{wh}-word \textit{WH} helps to build the set of alternatives in this system, by supplying alternatives that could fill a position in a complex numeral.

4.2. Pre-numeral \textit{some} and pragmatic halos

The pre-numeral \textit{some} has a different interpretation from the post-numeral interpretation, namely in having an approximative rather than “at least” interpretation. Whereas \textit{twenty-some} has an interpretation where any number from the range 21 to 29 would satisfy the phrase, \textit{some twenty} requires numbers close to 20, such as 18, 19, or 21. The numbers implied in \textit{some twenty} do not have to have 20 as their lower bound; they can start below 20 as well. Since pre-numeral \textit{some} doesn’t depend on the syntactic form of the numeral, I will assume that there is a different mode of approximation at work in the pre-numeral construction, and that the covert \textit{wh}-word implicated in the post-numeral construction is not used in the pre-numeral construction.

The interpretation in the pre-numeral \textit{some} cases seems most closely related to imprecision (Lasersohn, 1999; Kennedy, 2007). The way I will model this is by appealing to Lasersohn’s pragmatic halos. Lasersohn offers halos as an
explanation for imprecision, where natural language expressions have some amount of fuzziness surrounding them about what counts for an expression in a context. For example, *three o’clock* can be used imprecisely to mean 2:58pm in many contexts, due to *three o’clock* having 2:58pm within its pragmatic halo. As suggested by Morzycki (2011), halos might play a role in the compositional semantics, where he formalizes them using an alternative semantics. For my purposes here, the issue is how to get a halo around the number in the first place. I propose that the halo is coerced via presupposition accommodation, namely to satisfy the felicity requirements of *some*.

Well-known is that the determiner *some* enforces epistemic requirements on the speaker, namely that the referent of the *some* indefinite be unidentified. Strawson (1974) observes that this contrasts with *a(n)* indefinites, which do not have the same requirement.

\[(30)\]
\[
a. I’ve been stung by a wasp.
b. #I’ve been stung by some wasp.
\]

Strawson argues that (30b) is odd because of the felicity requirements of *some*. Wasps are normally not individually identifiable to the average person. Uttering the sentence generates the implication that the speaker could have in principle identified the wasp, but our own knowledge tells us that wasps cannot be identified. The tension between our knowledge of wasps and the implicature generated by the sentence causes us to judge the sentence as being odd.

To generate the unidentifiability requirement of *some*, Weir (2012) proposes that *some* incorporates an anti-singleton presupposition on its domain. This follows Alonso-Ovalle and Menéndez-Benito (2010), who originally propose a similar requirement on Spanish *algún*. (31) demonstrates this (*f* is a subset selection function).

\[(31) \{algún\} = \lambda f_{et,et} \lambda P_{et} \lambda Q_{et} : \text{anti-singleton}(f). \exists x [ f(P)(x) \land Q(x) ] \]

(Alongo-Ovalle and Menéndez-Benito 2010 : 19)

The anti-singleton presupposition is intended to generate an implication that the speaker cannot or will not identify the referent of the indefinite noun phrase.

\[(32) \text{Choice Functional Some (Final Version)} \]
\[
[\text{some } \alpha] = [\alpha] \text{ is not a singleton . } \{ f([\alpha]) \}
\]

where *f* is a choice functional variable

The presupposition is satisfied in the post-numeral *some* case, due to fact that the covert *wh*-word supplies a set of alternatives for *some* to choose from. In the pre-numeral case, however, there is no non-singleton set of alternatives, since numerals denote singletons. The anti-singleton presupposition fixes this
problem; the presupposition is accommodated by assuming that the number that some combines with does in fact denote a non-singleton. The mechanism to do this is to union the denotation of twenty with its pragmatic halo (schematically as in (34), where halo is a contextually sensitive function returning the pragmatic halo of some linguistic object).

(33) \([\text{some twenty}] = [\text{twenty}] \text{is not a singleton} \cdot f([\text{twenty}])\)  
Presupposition failure!
(34) \([\text{twenty}] = [\text{twenty}] \cup \text{halo}_c([\text{twenty}])\)

The lesson is that the pragmatic halo can be present just when we need it; it’s accommodated due to the pragmatic requirements of some.

5. Conclusion

In this paper I show that there are two approximative constructions using some with separate semantic representations, but that they can be treated in similar ways by making use of a choice functional analysis of some, and by making alternatives available in the semantics. Theoretically interesting in this analysis is the source of the alternatives. In the post-numeral some construction, the alternatives are generated through merger of a covert wh-word. The wh-word is interpreted in situ, where it directly denotes a set of numerical alternatives that are possible in the syntactic position on the wh-word. These alternatives are determined by the syntactic environment of the wh-word, making the post-numeral some sensitive to the syntactic properties of the numeral it combines with. The alternatives in the pre-numeral some, instead, are coerced to match the anti-singleton requirement of some; the pragmatic halo of the numeral is used for the set of alternatives in this case.

Notes

I thank Marcin Morzycki, Ai Taniguchi, Ai Kubota, Yusuke Kubota, Adam Gobeski, Taehoon Kim, Adina Williams, the audience at WECOL/AZLS 2013, and the Michigan State University semanticists for their comments and suggestions. All errors are my own.

References

Karttunen, Lauri. 1977. “Syntax and semantics of questions”. Linguistics and Philosophy, 1.1: 3-44.