This printout has been approved by me, the author. Any mistakes in this printout will not be fixed by the publisher. Here is my signature and the date: _____

# The role of semantics in licensing English synthetic compounds

**Andrew McKenzie**

## 1. Introduction

This paper argues that synthetic compounds in English (like *truck driver* or *car-chasing*) are licensed by the semantics rather than the syntax. They are available when an operator in the derived form binds the entity argument in the noun's denotation.

A longstanding problem for generative word-building involves English synthetic compounds (SC), which are grammatical but built from impossible verbs.

(1) *truck driver* > **Alan* ***truck-drives*** *for a living.*

(2) *car-chasing* (dog) > **The dog* ***car-chased*** *down the street.*

Previous approaches find that the syntax can build these verbs, but argue that another module blocks them from spell-out or use. For instance, Harley (2011) argues that prosody prevents incorporating verbalizing head $v^\circ$ from spelling out complements with multiple roots.

(3)


This approach is argued against by McIntyre (2014), who points out that zero-conversion routinely spells out such heads, in words like *grandstand*, *bear hug*, or *highlight*. I will show that semantic facts matter most, and argue that the module blocking unavailable SCs is the semantics.

## 2. Semantic facts matter most

In this section, I will point out three crucial semantic facts about the formation of noun-verb SCs. First, the ungrammatical examples tend to involve the incorporation of a noun to its verb.[1] However, non-object thematic roles are routinely grammatical and highly productive.

(4) *I* ***hand-wash*** *every fork.* (instrument) (Pullum & Huddleston, 2002)

(5) *The kid* ***daydreams*** *constantly.* (time)

(6) *The mayor can* ***speed-read*** *like you wouldn't believe.* (usage)

Second, the noun is often interpreted like a *de dicto* or intensional noun, because it describes entities that may not exist in the actual world, because their interpretation is subordinated to a modal.

* Andrew McKenzie, University of Kansas, andrew.mckenzie@ku.edu. Thanks go to my Kiowa consultants: Dolores Harragarra, Dorothy Delaune, Juanita Ahtone, and the late Christina Simmons. Also to Marianne McKenzie, Gus Palmer, Jr., and the audiences in the Research in Field Languages group at the University of Kansas. Part of this research was funded by NSF/NEH grant #BCS-1664431

[1] I use *incorporation* in a loose sense for nomenclature; nothing hinges on the use of the term.


Cascadilla Proceedings Project
Completed July 3, 2018

(7) *Terese **pimp-slapped** the mayor.*
⇝ slap in the ostentatious manner of a (generic/stereotypical) pimp

(8) *The **dog catcher** has never had to catch any dogs.*
⇝ dogs who they would catch should the need arise

Third, object-incorporation is allowed in lexicalized exceptions. These exceptions tend to be 'nameworthy' (Mithun, 1984) in a culture or subculture. For instance, the verb in (9) is used in rural subcultures, while that in (10) is used in science ficition or fantasy writing to describe the creation of a fictional universe.

(9) *We **deer hunt** every year.*

(10) *Pratchett **world-builds** like no other.*

Fourth, whatever these semantic contexts are doing to license SCs must occur locally, within the word, because inflectional and aspectual environments don't license SC— like the generic in (1).

(1) **Alex **truck-drives** for a living*.

## 3. The role of semantics in licensing

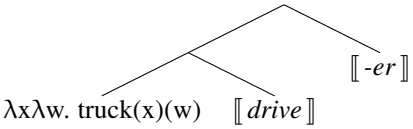
These facts are linked if the licensing of SCs depends on the semantics. The syntax builds these compounds grammatically, but they are only interpreted if an operator binds the noun's entity argument. The operator can be provided by a number of sources, notably derivational affixes and null thematic linkers.

### *3.1. Provided by derivational affix*

This is the case for synthetic compounds, like *truck+drive-er*. The noun composes semantically with the verb by conjunction (via Intensional Restrict[2]), and the affix takes the result as its argument. It quantifies over the entity, event, and world arguments of the compound verb. The world binder is essentially a modal, whose modal base MB(w) returns a set of worlds that allows for intensional *de dicto* interpretations.

(11) $[\![ \textit{truck driver} ]\!] \simeq \lambda x \lambda w.$ in worlds where x does their job, there are trucks and events of x driving them

$$\lambda x \lambda w.\ \forall w'[\ MB(w)(w') \rightarrow \exists y[\ truck(y)(w') \ \&\ \exists e[\ agent(x)(e)(w') \ \&\ drive(y)(e)(w')\ ]]]$$




### *3.2. Provided by null thematic linker*

Non-object nouns can freely combine with verbs because they need a thematic role, and a binder provides it. For instance, *hand-wash* does not mean 'wash the hands' but 'wash by hand', requiring a thematic role of instrument to be applied somehow. Assuming strong compositionality, some kind of null head ($f_{inv}$) provides this role for anything but themes, which are provided by the verb itself (Kratzer, 1996) and agents, which are ruled out by standard constraints on incorporation.

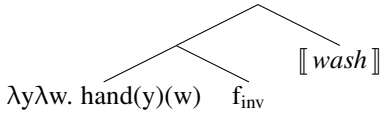
For the sake of convenience, I will employ a vague thematic role Involvement, which holds of an entity that is obliquely involved in an event. The nature of the involvement can be left to pragmatics. The head $f_{inv}$ takes the noun as its argument (12) and combines by Event Identification with the verb.

[2] Strictly speaking, mathematical constraints make Restrict impossible in a possible-world semantics. The problem is obviated with situation semantics, so we will not dwell upon it here.

(12) $f_{inv} = \lambda P_{e,t} \lambda e \lambda w.\ \exists y[\ P(y)(w)\ \&\ Involve(y)(e)(w)\ ]$

(13) $[\![\ \textit{hand-wash}\ ]\!] \simeq \lambda x \lambda e \lambda w.$ e is an event of washing x in w and there are hands y that are obliquely involved in e in w.

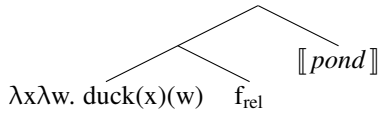
$\lambda x \lambda e \lambda w.\ wash(x)(e)(w)\ \&\ \exists y[\ hand(y)(w)\ \&\ Involve(y)(e)(w)\ ]$




The head is located between the verb and the noun, to allow for a thematic role to be assigned and linked to the verb. A similar kind of linker (call it $f_{rel}$) is also seen in ordinary compounds. Located between the two stems, $f_{rel}$ obviates the weak compositionality in their combined meaning.

(14) *duck pond* $\simeq \lambda x \lambda w.$ x is a pond in w and in worlds where x's purpose is fulfilled, there are ducks who use x

$\lambda x \lambda w.\ pond(x)(w)\ \&\ \forall w'[\ MB(w)(w') \rightarrow$
$\exists y[\ duck(y)(w')\ \&\ \exists e[\ agent(y)(e)(w')\ \&\ use(x)(e)(w')\ ]]]$

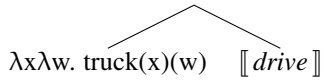



## 4. Blocking without the semantic licenser

Without a binding operator, the result is either incomplete and uninterpretable, or leads to an unattested meaning. In (15), the lack of an affix leaves the compound verb with an unsaturated entity argument.

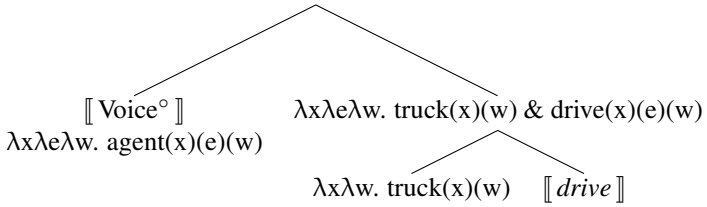
(15) No affix : **truck-drive* (via Intensional Restrict)

$\lambda x \lambda e \lambda w.\ truck(x)(w)\ \&\ drive(x)(e)(w)$




This operator must come below Voice° lest the agent be the object—*truck-drive* would be inherently reflexive.

(16) $\lambda x \lambda e \lambda w.\ agent(x)(e)(w)\ \&\ truck(x)(w)\ \&\ drive(x)(e)(w)$

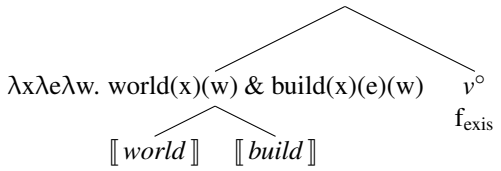



One can predict that object SCs should be licensed when they are reflexive, and *self-drive* is a verb.

If the agent-introducing operator somehow avoided combining the agent and object, or if we eschewed Neo-Davidsonian semantics altogether, the expression would remain forever uninterpretable, because nothing higher in the structure would saturate the object's entity argument.

What about nameworthy exceptions? These, along with the affixation cases, signal a role for derivation in licensing SCs. Specifcially, the lexicalization process somehow allows the introduction of a binder above the verb+noun compound, in this case an existential operator $f_{exis}$. This might be housed in the categorizing head $v^{\circ}$.

(17) ⟦ *world-build* ⟧ ≃ λeλw. there is a fictional world x and e is an event in w of creating x




One question raised by this approach is why this process is unavailable in productive verb-building. A preliminary answer suggests a distinction between purely syntactic $v^{\circ}$ heads and derivational ones.

## 5. A prediction supported by Kiowa incorporation

If noun incorporation involves binding operators, we ought to see unexpected licensing of object noun combinations when binding operators are used. Evidence supporting this prediction comes from the Kiowa language, spoken in the central US state of Oklahoma.[3] Kiowa allows noun incorporation of obliques but generally bars object incorporation (18). For instance, (19a) can mean 'he ate it like butter', but elicitation finds that it cannot mean 'he ate butter' (19b).

(18) * *Belle àn ∅-**kấu**+àumàu.*
B. HAB 3S>3S-shawl+make.IPFV
'Belle makes shawls.' (Adger et al., 2009) [*kấu* 'shawl' triggers singular agreement]

(19) a. *∅-**báuláu**+fàulè.*
3S>3S-butter+eat.IPFV.EVID
'He ate it like butter.' (Watkins, 1990)
b. ⟦ *∅-**báuláu**+fàulè* ⟧ ≠ 'he ate butter' (field notes)

As in other languages, lexicalized 'nameworthy' exceptions like (20) allow for object incorporation. Many of these exceptions have since been made into idioms.

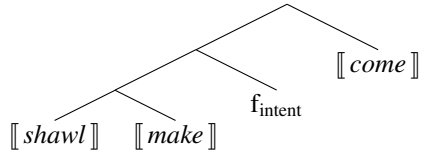
(20) *àn é-**qī**+gũgù.*
HAB 3S>3INV-firewood+put.IPFV
'She lights campfires.' (f.n.) [INV = inverse number, here singular]

Interestingly, there exists one context where Kiowa productively allows object incorporation: If the noun is the object of a verb that is itself incorporated. Speakers will happily accept (21) an instant after swiftly rejecting (18).

(21) *Belle ∅-**kấu**+àum+chàn.*
B. 3S-shawl+make+come.PFV
'Belle came to make shawls.' (f.n.)

Kiowa uses verb incorporation for some intensional contexts, like the intent/control context in (21). These embedded verbs require a binder over their event and world arguments ($f_{intent}$ in this case), since those differ from those of the main verb. Since that binder is already required, it is a small matter for it to also bind the verb's entity argument, which in turn binds the noun's. This licenses object incorporation.

(22) λxλeλw. come(x)(e)(w) & ∀w′ where x's intent for e comes true,
∃y[ shawl(y)(w′) & ∃e′[ make(y)(e)(w′) & agent(x)(e)(w′) ]]




[3] ISO-639-3: kio, Endonym *Cáuijọ̄gà* [kɔ́j.tɔ̃ːʷ.ɟæ̀]. Kiowa-Tanoan group. No standard orthography, but examples are written in the popular Parker McKenzie orthography. Glossing is standard Leipzig rules, with INV for inverse number marking.

## 6. Conclusion

This paper has demonstrated that English synthetic compounds are made fully interpretable by their affix, which binds the noun's entity argument. Meanwhile, the verbs that they are built from lack such a binder, and are uninterpretable unless a special context provides some other binder.

## References


Adger, David, Daniel Harbour & Laurel Watkins (2009). *Mirrors and Microparameters: Phrase Structure Beyond Free Word Order*. Cambridge Univ. Press.

Harley, Heidi (2011). *Compounding in Distributed Morphology*. Oxford University Press, Oxford.

Kratzer, Angelika (1996). Severing the external argument from its verb. Johan Rooryck, Laurie Zaring (ed.), *Phrase Structure and the Lexicon*, Kluwer, Dordrecht, 109–137.

McIntyre, Andrew (2014). Synthetic compounds and argument structure: Messages from a *bandwagon-jumper-onner* and a *two-cents-worth-thrower-inner*, URL `http://bit.ly/2hIrRle`. Available online.

Mithun, Marianne (1984). The evolution of noun incorporation. *Language* 60:4, 847–894.

Pullum, Geoffrey K. & Rodney Huddleston (2002). *The Cambridge Grammar of the English Language*. Cambridge University Press.

Watkins, Laurel J. (1990). Noun phrase versus zero in Kiowa discourse. *IJAL* 56:3, 410–426.