Theme vowels are verbs

Antonio Fábregas

University of Tromsø

1   Introduction and main claims

The goal of this squib is to provide several empirical arguments that support the view that the real verbal predicate in Spanish are the theme vowels that appear combined with what is traditional regarded as the lexical verb (1). In a nutshell, then, theme vowels are light verbs,[1] as suggested by Kayne (2016) following Solà (1994).

(1)   a.   cant-**a**
              sing-THV1
              'sing'
       b.   beb-**e**
              drink-THV2
              'drink'
       c.   viv-**i**
              live-THV3
              'live'

While in examples such as (1) this claim is essentially trivial if one assumes that the unit that the theme vowel combines with is an acategorial root (Marantz 1997; Arad 2003; Marvin 2002), the apparent problem of this approach is that theme vowels also seem to appear with morphological verbalisers, as in the example (2). The standard analysis of such cases in Neo-Constructionist approaches is that the first constituent is a root, while the verbaliser should be decomposed in two parts: *ific* as the spell out of the verbal layer, and *a* as the theme vowel.

---

[1]Kayne himself admits that the notion of 'light verb' is difficult to define in pretheoretical terms, and it is largely an empirical issue to determine which properties it should have. Here we adopt as a working definition the idea that a light verb should be viewed as a constituent that licenses the presence of verbal functional structure, such as tense, aspect, mood and subject or object agreement, but does not introduce enough conceptual information to be a predicate by itself; the light can also provide syntactic positions for arguments, but does not determine autonomously the interpretation of those arguments. Some verbs are always light in this sense, such as *ser* and *estar*, while some full verbs can double as light verbs, as *poner*. See Butt (1995; 2003); Butt & Geuder (2001); Grimshaw & Mester (1988); Lin (2001); Mohammad & Karimi (1992) for further properties.

(2) clas- ifica
     class- ify
     'classify'

We will also provide arguments that in (2) the verbaliser is in fact the theme vowel, and the *ific* constituent should be treated as a root modifying it (Lowenstamm 2014). Thus, we will argue for the following identity:

(3) Theme vowel = Light verb = Verbaliser

The relevance of this empirical observation is that it makes it easier to dissolve a frequently cited *prima facie* counterargument against syntactic approaches to word formation. Blevins (2007), for instance, points out that the fact that some languages have theme vowels while others don't supports an idiosyncratic—and therefore lexicalist—treatment of word formation to the extent that whether a language has or does not have theme vowels is an arbitrary property of how verbs acquire their morphological shape with no consequences for syntax or semantics (unlike for instance PRO DROP, V2 or RICH AGREEMENT). If theme vowels are light verbs, then all languages have 'theme vowels' in the deep sense because all languages have verbs (cf. Kayne 2016 for the claim that English *e* in *grad-e-d* is a theme vowel). Second, theme vowels do not seem to correspond to independent syntactic objects in the available theories. In lexicalist proposals, the base is already a verb, and the theme vowel just marks the conjugation class (4). In Neo-constructionist theories (Halle & Marantz 1993), the verbaliser is the real verb and the theme vowel has to be introduced post-syntactically (as a dissociated morpheme, cf. Oltra-Massuet 1999) to mark the conjugation class (5).

(4) [$_V$ cant-a]

(5) [[[$_\sqrt{}$clas] ific]$_v$ + a]

This second problem is also dissolved, because if this proposal is right the theme vowel would be the spell out of the verbal layer(s), a relatively well-understood syntactic and semantic entity.

2   Theme vowels = light verbs

Kayne (2016) argues, following an initial suggestion by Solà (1994), that the right segmentation of an English regular past tense is (6).

(6) rain$^{\text{Root}}$-e$^{\text{THV}}$-d$^{\text{Tense}}$

If theme vowels are light verbs, then a theme vowel and a light verb would in principle be introduced in the same position—assuming for the time being that there is a unique syntactic position for both. From here it follows that the best candidates to be light verbs in English should lack any presence of the /e/ that corresponds to the theme vowel, a prediction that seems borne out:

(7)   *beed, *haved, *doed, *goed, *comed, *taked, *bringed, *putted, *getted, *gived, *maked, *letted, *sayed, *canned, *willed, *shalled, *mayed, *musted

In what follows I will extend Kayne's observation to Spanish. Let us take the prototypical case of a light verb in Spanish, *ser*. As can be seen in (8), this verb is prototypically light in at least two senses: it lacks enough conceptual semantics to stand alone as a predicate, as witnessed by (8b), and it has plain uses as an auxiliary, in particular as a passive auxiliary (8c).

(8)   a.   Juan es alto.
             Juan is tall
             'Juan is tall'
       b.   *Juan es.
             Juan is
             'Juan is.'
       c.   El prisionero fue detenido.
             the prisoner   was detained
             'The prisoner was detained.'

In the present paradigm it is impossible to identify any segment that could plausibly correspond to the THV.

(9)   SG   1   so-y
             2   ere-s
             3   es
       PL   1   so-mos
             2   so-is
             3   so-n

Moreover there is no evidence that here we have a root that combines with the theme vowel. The few cases that can be argued to relate to this verb through word formation are bizarre, infinitives or highly lexicalised expressions that are quite likely to come from Latin.

(10)   a.   esencia
              is.ance
              'essence' (not 'state of being')

3

b. ser
to.be
'entity'

My claim is that both facts—absence of a root+theme vowel make up and its light verb nature—are correlated and *be* lacks a theme vowel because both *be* and the theme vowel are the same object in the syntax. This explains the facts if we have an alternation along the shape of (11), with a 'standard' verb consisting of the light verb and the root, and *ser* being just the light verb without any root:

(11)   VP
         V   √
         a   cant-

(12)   VP
        SER
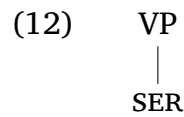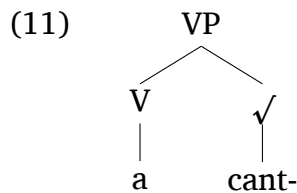
Another prototypical case of light verbs are those used as aspectual auxiliaries: the progressive *estar* 'be', the perfect *haber* 'have' and the prospective *ir* 'go'.

(13)   a.   Juan está comiendo.
              Juan is   eating
              'Juan is eating.'
    b.   Juan ha comido.
              Juan has eaten
              'Juan has eaten.'
    c.   Juan va   a comer
              Juan goes to eat
              'Juan is going to eat.'

The first two do not have a use as semantically 'strong' verbs, but they are closely related in existential sentences, where they contrast depending on the definiteness of the only argument, with definites combining with *estar* (14b).

(14)   a.   Hay   un libro.
              there.is a   book
              'There is a book.'
    b.   Estaba   el libro.
              there.was the book
              'The book was there.'

The third verb is closely related to *ser* in its morphology: the perfective form of *ir* and *ser* is morphologically identical.

(15)   fui (3SG.PRF of *be* and *go*)

I want to argue that these three light verbs also lack a theme vowel in the present. Consider their paradigms.

(16)

		estar	*haber*	*ir*
SG	1	esto-y	he	vo-y
	2	está-s	ha-s	va-s
	3	está	ha/ha-y	va
PL	1	esta-mos	he-mos	va-mos
	2	está-is	habé-is	va-is
	3	está-n	ha-n	va-n

Prima facie, the verb *estar* could be interpreted as consisting of a root *est-* and a first conjugation theme vowel *a,* but this causes problems for the algorythm that assigns stress in Spanish according to Oltra-Massuet & Arregi (2005), which should assign stress to the first vowel, as in *cantar.* The problem disappears if the verb is segmented as (19).

(17)   a.   /es.tá/ (cf. /kán.ta/)
          b.   /es.tán/ (cf. /kán.tan/)

(18)   a.   est-á
          b.   est-á-n

(19)   a.   está
          b.   está-n

In the case of *ir,* there is simply no evidence that *va* should be segmented: *v-* is not used as a root. In the case of *haber,* with the only possible exception of the 2pl *habéis,* which could be amenable to a segmentation *hab-é-is,* there is no segment that could correspond to the root—remember that in Spanish the letter 'h' has no sound. The same claim should be extended to the verb *dar,* another strong candidate to being a light verb:

(20)   a.   doy
          b.   das
          c.   da
          d.   damos
          e.   dais
          f.   dan

Here are some pieces of additional evidence that these verbs lack any theme vowel.

First of all, we have the fact that these verbs do not make an indefinite past as one would expect if the *a* were a Theme Vowel. The indefinite past of *cant-a* is *cant-a-ste,* but these verbs do not follow this pattern:

(21)   a.   *estaste
       b.   estuviste

(22)   a.   *daste
       b.   diste

(23)   a.   *vaste
       b.   fuiste

Second, nobody, to the best of my knowledge, has ever tried to regularise the past tense of these verbs, unlike what is the case with the verb *andar* (24), which can be segmented in a theme vowel and a root (25).[2]

(24)   a.   and-a-ste
       b.   anduv-i-ste

(25)   a.   /án.da/
       b.   /án.dan/

Contrast this with our light verbs that include the sequence */ub/* in their indefinite past:

(26)   a.   *estiste / estuviste
       b.   *histe / hubiste

Another interesting generalisation regarding these verbs without a theme vowel—albeit one I have no explanation for—is that they are precisely those that add a segment /I/ (ortographically, *-y*) in one form of the present indicative paradigm, normally the 1SG (with the verb *haber* adding it only in presentational contexts):

(27)   a.   so-y, esto-y, vo-y, do-y
       b.   ha-y

At the very least, that these verbs—and only these verbs—take this morpheme strongly suggests that they share some property with each other. My suggestion is that the property is that they are, in themselves, theme vowels.

---

[2]As far as we can tell, the alternation between *andaste* and *anduviste* does not reflect two uses of the same verb, one as full verb and one as light verb. The form *anduviste* is getting lost in contemporary Spanish, and it is safe to say that speakers that use it learn it at schools, and some of them quickly forget them. For this reason, pending further research, it seems plausible to us that speakers treat *anduv* as an allomorph of the root and the structure of this verb is always the one associated to full verbs.

3   Verbaliser = Theme Vowel = Light verb

There is a long tradition that views the verbalisers as light verbs (Harley 1995). Just as what we have descriptively labeled theme vowels, they combine with roots, and are associated to abstract meanings which in compositional cases function as semantic skeletons. Treating the verbalisers as light verbs is therefore not an unheard of idea, but the problem emerges when we combine this with the proposal that the theme vowel itself is a light verb, essentially because in principle the verbaliser and the theme vowel seem to co-occur. (28) lists the most productive verbalisers in contemporary Spanish; note that the last segment in each one of them is morphophonologically identical to a theme vowel, and also defines the resulting complex verb as belonging to the conjugation class associated to that theme vowel.

(28)   a. iz**a**   (autor-iza 'authorise')
        b. ific**a** (clas-ifica 'classify')
        c. it**a**   (debil-ita 'debilitate')
        d. ec**e**   (palid-ece 'to become pale')
        e. e**a**   (tont-ea 'to act silly')

Thus, a verb like *palidecer* inflects as a second conjugation verb, as expected if the final *e* in *ece* is the second conjugation theme vowel (29). A verb like *autorizar* inflects in the first conjugation (30), again as expected if the final segment is the theme vowel.

(29)   a. comer > com-í-a (IPFV, 3SG)
        b. palidecer > palid-ecí-a (IPFV, 3SG)

(30)   a. cantar > cant-a-ba (IPFV, 3SG)
        b. autorizar > autor-iza-ba (IPFV, 3SG)

I argue that in fact these verbalisers are allomorphs of the theme vowels.[3]

I can provide three arguments supporting this analysis. First of all, the traditionally considered 'verbalisers' *ific, it, ec* and so on never ever appear without the segment we believe is the theme vowel. There are no plausible phonological reasons for this. Removing the theme vowel from *ifica* before a nominaliser like *ción* 'ation' would produce a sequence that

---

[3]An alternative would be to say that *ific, ec, iz* are roots modifying the light verb, as argued by Lowenstamm (2014) in the case of some apparent adjectivalisers / nominalisers. There are two reasons not to adopt this analysis: first, I am not aware of any use of the relevant morphemes as category-changing affixes producing nominalisations or adjectivalisations, which would make surprising that they are roots. Second, as Tarald Taraldsen (p.c.) made us notice, if they were modifiers of the VP layer, then we would expect them to be able to combine with light verbs.

is perfectly possible in other contexts, such as (31).

(31)   destruc-ción

Still there are no words in Spanish ending in a sequence *ificción* which can be plausibly analysed as the nominalisation of a verb. The same applies to **itción, *eción, *izción*. Note in contrast that there are many words ending in *-facción*:

(32)   a.   calefacción (heating)
         b.   rarefacción (rarefaction)
         c.   licuefacción (liquefaction)
         d.   putrefacción (putrefaction)

Importantly, in none of these cases does the noun derive from a verb:

(33)   a.   *calefacer
         b.   *rarefacer
         c.   *licuefacer
         d.   *putrefacer

The only case to my mind where there is a sequence *facción* and there is a related verb is *satisfacer* 'satisfy', where in case one wanted to segment it, the base would be *hacer* and the first member would be a prefix, not a verbalised root.

This property of verbalisers is explained automatically if *ifica* is just a version of the theme vowel. An account where *ific* is the verbaliser itself has no way of blocking that the theme vowel, itself a distinct morpheme, can be absent.

A second argument in support of the view that the verbalisers are the theme vowels is that they never appear in light verbs. None of these affixes can attach to something to make a light verb, something that we expect given that light verbs occupy the position of the theme vowel.

The third argument is that the connection between the morphemes *ific, e, it, ec* and the theme vowel is univocal. It is never the case that, for instance, *ific* appears with the theme vowel *e* or *i* instead of *a* (34).

(34)   a.   *ificer
         b.   *icer
         c.   *eir
         d.   *itir

This is not logically necessary, given that for instance with nominalisers and adjectivalisers, Spanish allows that the morpheme responsible for the category assignment combines with more than one distinct marking. In (35), the nominaliser *-ez* combines with *a* marking feminine in the

first case, but with null marking in the second.

(35)   a.   pobr-ez-a
              poor-ness-A
              'poverty'
        b.   tont-ez-ø
              silly-ness-ø
              'stupidity'

The pattern is however explained if in the case of theme vowels they are the real verbalisers, and an exponent like *ifica* is another spell out they can get. In the case of nominalisers and adjectivalisers, we assume—following the traditional approach—that they are themselves responsible for the category change and the final vowel is a higher functional head. This also explains that in the case of nominalisers and adjectivalisers this final vowel is systematically cancelled:

(36)   a.   ceremoni- os-   o
              ceremoni- ous- O
              'ceremonious'
        b.   ceremoni- os-   idad
              ceremoni- ous- ity
              'ceremoniosity'
        c.   *ceremoni- os-   o- idad
              ceremoni- ous- O- ity
              Intended: 'ceremoniosity'

Note, incidentally, that *ifica* can be historically related to the light verb *facere* 'to make / do' in Latin (cf. Spanish *hacer*).

4   Conclusions and further prospects

There are, therefore, several empirical arguments that there should be total structural identity between theme vowels, verbalisers and light verbs in the sense that they identify the same position within the verbal structure. If these arguments are right, then we would be one step closer to reducing what seems to be surface, idiosyncratic morphological variation to well-defined syntactic or semantic properties of universal structures.

Note that we have made the argument that light verbs are theme vowels—or vice versa—based on the present tense. One reason for this is that in the past tense additional segments that could be identified as theme vowels (Oltra-Massuet 1999) emerge:

(37)   a.   er-a
            was-THV
            'It was.'
       b.   esta-b-a
            be-IMPF-THV
            'It was.'

The way in which we interpret these facts follows Oltra-Massuet in the claim that the structure of the Romance verb includes more than one position for the theme vowel. Instead of treating theme vowels as dissociated morphemes, however, we will push our preliminary findings and propose that in (37) the visible theme vowel is a light predicate corresponding to [PST]. The question of how many theme vowels a particular verb form carries, in this view, is a function of the height at which the verb is initially introduced in the structure. A light verb like *ser* or *estar* is introduced below the position for Past Tense, and therefore it lexicalises a chunk of structure that does not include the projection that defines Past. Therefore, by the Exhaustive Lexicalisation Principle, when the light verb combines with the [PST] head, an additional exponent—itself a theme vowel—has to be combined with it. A natural extension of this proposal is to check whether, as expected, the number and position of the theme vowels that light predicates of different types carry follows the Functional Sequence, as one would expect from this type of treatment. We hope to explore this issue in the immediate future.

Acknowledgment

References

Arad, Maya. 2003. Locality constraints on the interpretation of roots: the case of Hebrew denominal verbs. *Natural Language and Linguistic Theory* 21. 737–778.

Blevins, James. 2007. Conjugation classes in Estonian. *Linguistica Uralica* XLIII(4). 250–267.

Butt, Miriam. 1995. *The structure of complex predicates in Urdu* (Dissertations in linguistics). Stanford, Ca.: CSLI Publications.

Butt, Miriam. 2003. The light verb jungle. In *Harvard working papers in linguistics*, vol. 9, 1–49. Cambridge, MA: Harvard.

Butt, Miriam & Wilhelm Geuder. 2001. On the (semi)lexical status of light verbs. In Norbert Corver & Henk van Riemsdijk (eds.), *Semi-lexical categories: the function of content words and the content of function words*, 000–000. Berlin: Mouton de Gruyter.

Grimshaw, Jane & Armin Mester. 1988. Light verbs and theta-marking. *Linguistic Inquiry* 19. 205–232.

Halle, Morris & Alec Marantz. 1993. Distributed Morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Harley, Heidi. 1995. *Subjects, events, and licensing*. Cambridge, MA: MIT dissertation.

Kayne, Richard. 2016. What is suppletive allomorphy? NYU.

Lin, Tzong-Hong. 2001. *Light verb syntax and the theory of phrase structure.* UC Irvine dissertation.

Lowenstamm, Jean. 2014. Derivational affixes as roots. In Artemis Alexiadou, Hagit Borer & Florian Schäfer (eds.), *The syntax of roots and the roots of syntax*, 230–259. Oxford: Oxford University Press.

Marantz, Alec. 1997. No escape from syntax: don't try morphological analysis in the privacy of your own lexicon. In Alexis Dimitriadis & Laura Siegel (eds.), *Proceedings of the 21st Annual Penn Linguistics Colloquium* (University of Pennsylvania Working Papers in Linguistics), 201–225. Philadelphia: University of Pennsylvania.

Marvin, Tatjana. 2002. *Topics in the stress and syntax of words*. Cambridge, MA: MIT dissertation.

Mohammad, Jan & Simin Karimi. 1992. Light verbs are taking over: complex verbs in Persian. In J.A. Nevis & V. Samiian (eds.), *Proceedings of wecol 5*, 195–212. Fresno: Fresno State University.

Oltra-Massuet, Isabel & Karlos Arregi. 2005. Stress-by-structure in Spanish. *Linguistic Inquiry* 36(1). 43–84.

Oltra-Massuet, Maria Isabel. 1999. *On the notion of theme vowel: a new approach to Catalan verbal morphology*. Cambridge, MA: Massachusetts Institute of Technology MA thesis.

Solà, Jaume. 1994. Morphology and word order in Germanic languages. In Werner Abraham (ed.), *Minimal ideas: syntactic studies in the minimalist framework*, 217–251. Amsterdam & Philadelphia: John Benjamins Publishing Co.