

The SAME constituent structure

Carlo Cecchetto

(University of Milan-Bicocca)

Alessandra Checchetto

(Ca' Foscari University, Venice)

Carlo Geraci

(Institut Jean-Nicod, Paris)

Sandro Zucchi

(University of Milan)

based on a joint work with Mirko Santoro

DO LANGUAGE UNIVERSALS EXIST?

Most formal linguists would answer “yes” to the question above. However, the idea that language universals exist has been recently attacked in the following paper, which explicitly tries to talk cognitive scientists out of the chomskyan paradigm:

Nicholas Evans & Stephen C. Levinson (2009), The myth of language universals: Language diversity and its importance for cognitive science, *Behavioral and Brain Sciences*, 32, 429–492 (including commentaries)

TWO QUOTATIONS

“....there are vanishingly few universals of language in the direct sense that all languages exhibit them. Instead, diversity can be found at almost every level of linguistic organization. This fundamentally changes the object of enquiry from a cognitive science perspective....”

Evans & Levinson (2009: 429)

“The cognitive sciences have been partially immunized against the proper consideration of language diversity by two tenets of Chomskyan origin. The first is that the differences are somehow superficial, and that expert linguistic eyes can spot the underlying common constructional bedrock. This, at first a working hypothesis, became a dogma, and it is wrong, in the straightforward sense that the experts either cannot formulate it clearly, or do not agree that it is true.”

Evans & Levinson (2009: 432)

CONSTITUENTS UNDER ATTACK

According to Evans & Levinson, a good example of a “common constructional bedrock” that “became a dogma” and “is wrong” is the notion of constituent. Since they explicitly say that sign languages are a special challenge to the universalist view they oppose, today we will take up this issue and investigate whether the notion of constituent is defensible and/or useful in sign languages.

We will show that building constituency tests for Italian Sign Language (LIS) is not only possible, but modality specific properties of sign languages can actually help fixing some problems that affect the reliability of these constituency tests in many spoken language.

Evans & Levinson are not very explicit about the notion they attack. They define constituency as follows:

“Constituency is the bracketing of elements (typically words) into higher-order elements”

Evans & Levinson (2009: 440)

In linguistics constituents are defined operatively by using a well-established set of tests. The most interesting ones are pro-form substitution and ellipsis, since they can reveal the underlying hierarchical structure of the sentence.

If these tests identify constituents in sign languages, we can conclude that sign languages have an underlying hierarchical structure. If we can show this, the property of being hierarchically organized (or structure dependence property, as Chomsky 1956 called it) would remain a good candidate for being a language universal, across spoken and sign languages alike.

THE CLAIM TO BE TESTED

Languages have a hierarchical organization, namely the clause contains units, called constituents, that in turn contain other units, and the decomposition of bigger constituents into smaller ones proceeds up (or down!) to the level where we find single words.

In a nutshell:

The structure of the clause is not flat, but hierarchical.

A classical way to show that the structure of the clause is not flat in English is applying tests like pro-form substitution or VP ellipsis:

- (1) John read a book in the dining room and Mario did so in the kitchen
- (2) John read a book in the dining room and Mario did so, too

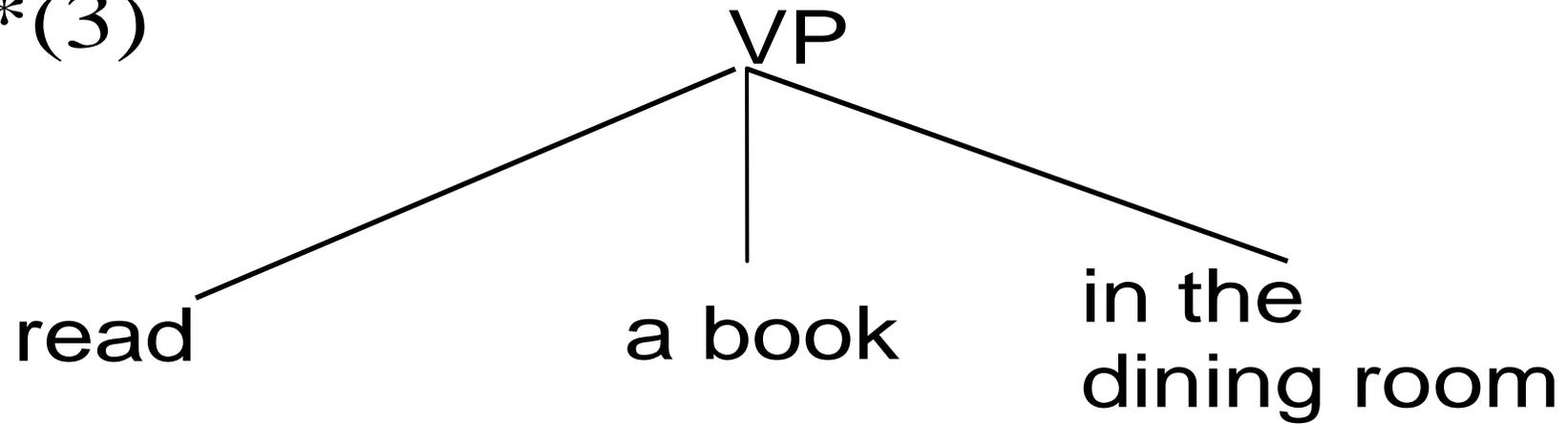
In (1) the missing constituent in the second clause is “read a book” while in (2) the missing constituent is “read a book in the dining room”.

(1) John read a book in the dining room and Mario did so in the kitchen

(2) John read a book in the dining room and Mario did so, too

These sentences suggests the verb phrase (VP) in the first clause in (1) has the hierarchical structure in (4) rather than the flat structure in (3). In particular, if the VP had the flat structure in (3), the grammaticality of (1) would be hard to explain, since there would be no node in the structure that can be omitted or replaced by a pro-form and contains “read a book” without containing “in the dining room”.

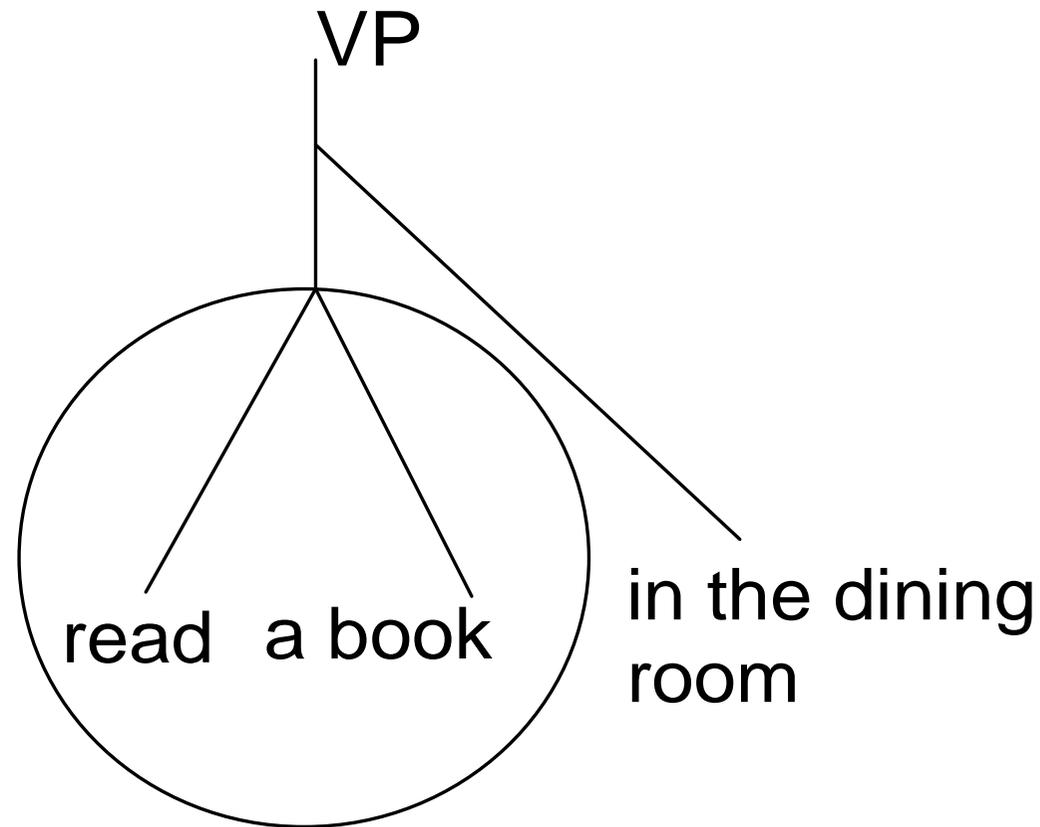
*(3)



Flat structure

* read + a book + in the dining room

√ (4)



Hierarchical structure

√ [[read a book] in the dining room]

(1) John read a book in the dining room and Mario did so in the kitchen

A MISCONCEPTION

Evans & Levinson tacitly assume that constituents are units of *contiguous* words. They have English in mind and English is a rigid word order language. Since many (probably most) languages have a freer word order than English, the notion of constituent would be hard to extend to these languages if constituents are defined as units of *adjacent words*.

However, by using tests like pro-form substitution and ellipsis it is possible to identify units of *discontinuous* words as constituents in languages other than English. The following example shows this for Italian:

(5) Gianni ha mangiato con piacere la minestra, invece Maria (l'ha fatto) controvoglia

Gianni has eaten gladly the soup , but Maria (it has done) unwillingly

In (5) the sequence of words [mangiato la minestra] is identified as a constituent although it is discontinuous (it is interrupted by the adverb “con piacere”).

THE HIERARCHICAL STRUCTURE OF THE VP IN LIS, INITIAL EVIDENCE

LIS has a construction in which a verb phrase can go un-uttered if a suitable antecedent is present. This construction involves the use of a sign glossed here as SAME (this sign is glossed as “stesso” in Italian). This construction is illustrated in (6) and (7):

(6) DINING-ROOM GIANNI VASE BREAK, MARIO SAME KITCHEN

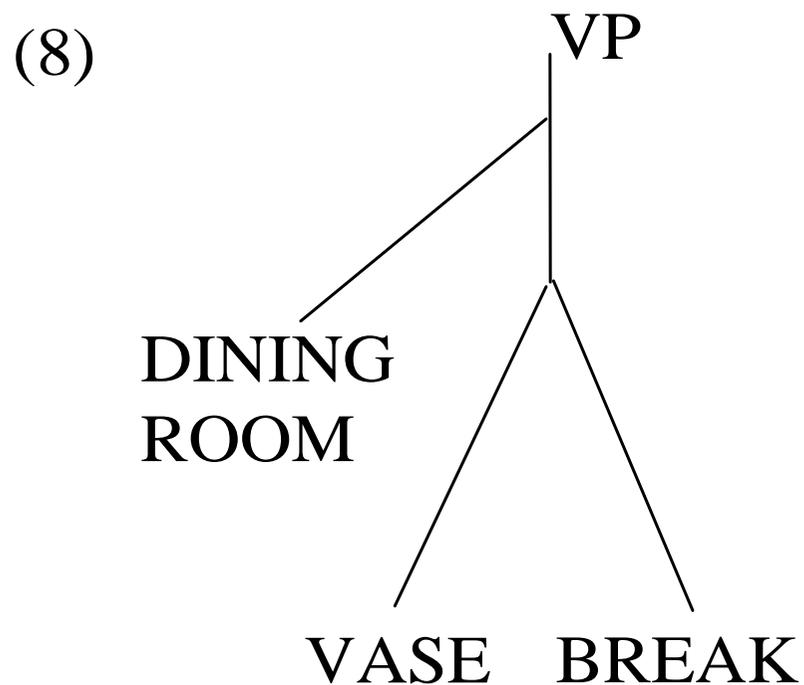
“Gianni broke a vase in the dining room and Mario did so in the kitchen”

(7) DINING-ROOM GIANNI VASE BREAK, MARIO SAME

“Gianni broke a vase in the dining room and Mario did so, too”

In (6) the VP VASE BREAK is missing, while in (7) the missing VP is DINING-ROOM VASE BREAK.

By the usual reasoning in order to explain the grammaticality of (6), we should assume that the VP in LIS is organized as in (8). In (8) we assume that LIS is head final.



NEGATION AND THE *SAME* CLAUSE

Another case that nicely illustrates the hierarchical structure of LIS is examples where the *SAME* clause is negated. For example, (9) is ambiguous:

(9) DINING-ROOM MARIA VASE BREAK PIERO *SAME* NOT

- (i) Maria broke a vase in the dining room, while Piero did not broke any vase
- (ii) Maria broke a vase in the dining room, while Piero broke a vase but not in the dining room

The meaning of (9) depends on the size of the VP which is negated in the *SAME* clause. If the constituent *VASE BREAK* is negated, the (i) reading arises. If the constituent *DINING-ROOM VASE BREAK* is negated, the (ii) reading arises.

(9) DINING-ROOM GIANNI VASE BREAK MARIO SAME NOT

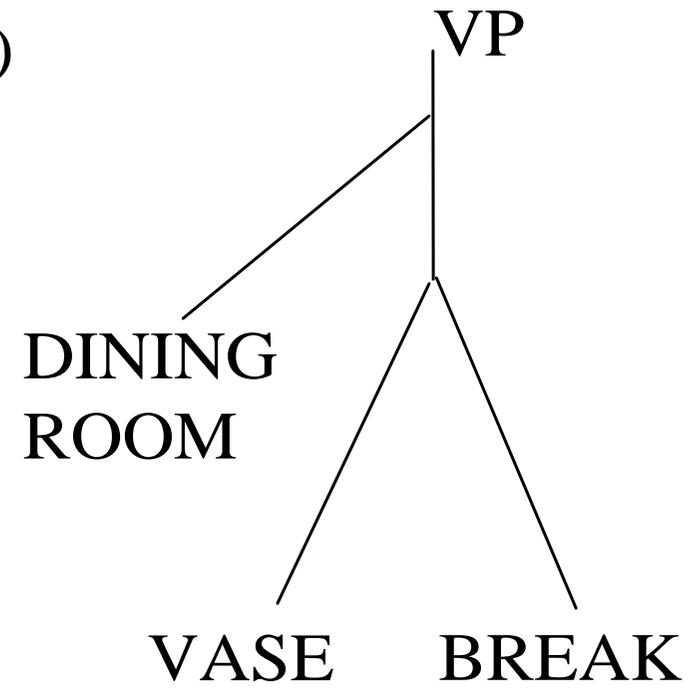
(i) Maria broke a vase in the dining room, while Piero did not break any vase

(ii) Maria broke a vase in the dining room, while Piero broke a vase but not in the dining room

Notice that the constituent identified by the reading in (ii) namely DINING-ROOM VASE BREAK is a discontinuous constituent, much like in the Italian case.

This tells us that at the relevant level of representation the VP has a structure like (10). We assume that the external argument is generated in the specifier position of vP and v takes the structure in (10) as its complement.

(10)



SAME AND AS-WELL

SAME is not the only sign that can accompany a missing VP in LIS. Another sign with a similar function is glossed AS-WELL (it is the sign glossed in Italian as “pure”).

An interesting difference between SAME and AS-WELL is that the latter seems to be a positive polarity item, as shown by the ungrammaticality of (12)

(11) DINING-ROOM GIANNI VASE BREAK MARIO AS-WELL

(12) *DINING-ROOM GIANNI VASE BREAK MARIO AS-WELL NOT

IS THIS TALK FINISHED?

So, the construction with SAME nicely reveals the hierarchal structure of the VP constituent in LIS. However, this might be too quick a conclusion.

I cannot finish my talk at this point (sorry about that!).

THE HIERARCHICAL STRUCTURE OF THE VP IN LIS, A POSSIBLE CONFOUNDING

One natural analysis is treating SAME as an adverbial expression much like “too” in English (cf. 13). Under this analysis, SAME co-occurs with VP ellipsis. According to an alternative analysis SAME is a verbal anaphora, much like “to do that” in (14).

(13) Gianni bought a new computer and Mario did too

(14) Gianni bought a new computer and Mario did that too

Deciding between the VP ellipsis and the verbal anaphora analysis is not without consequences. This is so because VP ellipsis is a more reliable test than pro-form substitution, at least when the missing constituent is complex. Indeed, verbal anaphors, at least in English, do not need to take linguistic antecedents, as shown by the classical examples in (15) from Hankamer & Sag (1976). The ungrammaticality of (15a) indicates that VP ellipsis requires a linguistic antecedent, while the grammaticality of (15b) indicates that a verbal anaphora does not require it (it suffices that the meaning of the anaphor is retrievable from the non-linguistic context).

[Hankamer attempts to stuff a 9-inch ball through a 6-inch hoop]

(15a) Sag: # It's not clear that you'll be able to.

(15b) Sag: ✓ It's not clear that you'll be able to do it.

Hankamer attempts to stuff a 9-inch ball through a 6-inch hoop]

(15a) Sag: # It's not clear that you'll be able to.

(15b) Sag: √ It's not clear that you'll be able to do it.

Therefore, if SAME is a verbal anaphor, one cannot exclude that the missing VP in the SAME clause is licensed by the non-linguistic context, much like “to do it” in (15b). If so, sentences with SAME do not tell us anything conclusive about constituent structure.

DECIDING BETWEEN THE VP ELLIPSIS AND THE VERBAL ANAPHORA ANALYSIS: CO-OCCURRENCE WITH A NON ELIDED VP

There are ways to set apart VP ellipsis and verbal anaphora. For example, an adverbial expression like “too” can accompany the non-elided VP (cf. 16) while a verbal anaphora compete with the non-elided VP for the same position, so the structure where a verbal anaphora and the corresponding full VP are both present is totally deviant, since they compete for the same position (cf. 17):

(16) Gianni bought a new computer and Mario bought a new computer too

(17) *Gianni bought a new computer and Mario did that bought a new computer too

So, if SAME can co-occur with a non elided VP, we have evidence for the VP ellipsis analysis.

DECIDING BETWEEN THE VP ELLIPSIS AND THE VERBAL ANAPHORA ANALYSIS: INCORPORATED ADVERBS

Another way to decide if a missing VP in the construction with SAME requires a linguistic antecedent or not relies on a modality specific feature of sign languages: in many sign languages, manner adverbs can either be incorporated in the verb (cf. 19a) or appear as an independent sign (cf. 19b). We illustrate this with LIS examples.

(19a) GIANNI EAT-QUICKLY

(19b) GIANNI EAT QUICKLY

Mario eats quickly

This allows us to check whether the missing VP requires a linguistic antecedent or can be licensed by the extra-linguistic context alone.

If the missing VP requires a linguistic antecedent, a sentence like (20) below should be fine, since a linguistic antecedent *is* available: this is the verb phrase MEAT EAT (the manner adverb QUICKLY is not an obligatory part of the antecedent VP). However, (21) should be ungrammatical since the adverb QUICKLY is incorporated into the verb, so the linguistic antecedent must be the verb phrase MEAT EAT-QUICKLY, but this triggers a clash in meaning (one cannot eat slowly and quickly at the same time).

(20) MARIO MEAT EAT QUICKLY. GIANNI SAME SLOWLY

(21) *MARIO MEAT EAT-QUICKLY. GIANNI SAME SLOWLY

(20) MARIO MEAT EAT QUICKLY. GIANNI SAME SLOWLY

(21) *MARIO MEAT EAT-QUICKLY. GIANNI SAME SLOWLY

On the other hand, if the meaning of the missing VP is retrieved from the extra-linguistic context, no contrast should arise between (20) and (21): the way in which the manner adverb is expressed in the first clause should not be important, as long as the activity of meat-eating is made contextually salient by the first sentence in the pair.

As a matter of fact, (21) is strongly ungrammatical.

So, this supports the hypothesis that the SAME construction involves VP ellipsis.

SHORT INTERLUDE, LESSONS FOR SPOKEN LANGUAGES COMING FROM SIGN LANGUAGES

The argument suggested by the interaction of adverb incorporation and use of SAME can be replicated in spoken languages like Italian:

(22) Gianni ha stramangiato la torta. *Piero l'ha fatto con moderazione
Gianni has over-eaten the cake. Piero did that moderately

(23) Gianni ha mangiato la torta con ingordigia. √ Piero l'ha fatto con moderazione

Gianni has eaten the cake greedily. Piero did that moderately.

In Italian the proform *farlo*, much like *to do that* in English, can take a non-linguistic antecedent. For example, if someone is jumping out of a skyscraper you can scream *non farlo!* (“don’t do that!”).

However, the ungrammaticality of (22) shows that *farlo* cannot take a non-linguistic antecedent in this sentence.

(22) Gianni ha stramangiato la torta. *Piero l’ha fatto con moderazione
Gianni has over-eaten the cake. Piero did that moderately

(23) Gianni ha mangiato la torta con ingordigia. ✓ Piero l’ha fatto con moderazione

Gianni has eaten the cake greedily. Piero did that moderately.

(22) Gianni ha stramangiato la torta. *Piero l'ha fatto con moderazione

Gianni has over-eaten the cake. Piero did that moderately

(23) Gianni ha mangiato la torta con ingordigia. √ Piero l'ha fatto con

moderazione

Gianni has eaten the cake greedily. Piero did that moderately.

The generalization emerging from data like (22) and (23) seems to be that even pro-forms that in principle can take a non-linguistic antecedent cannot do that if a linguistic antecedent is available in the relevant sentence. So, *farlo* in (22) must take as its antecedent *stramangiare la torta* but this creates a clash in meaning with the adverb 'moderately'.

DECIDING BETWEEN THE VP ELLIPSIS AND THE VERBAL ANAPHORA ANALYSIS: LICENSING UNDER PARALLELISM

(24) GIANNI BEANS EAT-QUICKLY

Sentences like (24) in which the adverb is incorporated into the verb are ambiguous. According to one reading, Gianni eats beans quickly (but he does not necessarily eat a big quantity of beans).

This is the “speed reading”.

(24) GIANNI BEANS EAT-QUICKLY.

According to the second reading of sentence (24), Gianni eats a lot of beans.

This is the “amount reading”.

If the adverb is not incorporated, only the “speed reading” is selected.

This ambiguity allows us to observe a strong parallelism requirement between the antecedent clause and the SAME clause, namely the reading that obtains in the antecedent clause must obligatorily obtain in the SAME clause as well:

(25) GIANNI BEANS EAT-QUICKLY. PIERO SAME

If the “speed reading” is selected in the antecedent clause, the SAME clause receives the “speed reading” (Piero eats quickly).

If the “amount reading” is selected in the antecedent clause, the SAME clause receives the “amount reading” (Piero eats a lot of beans).

This parallelism condition is consistent with the hypothesis that the SAME clause involves VP ellipsis, since it is well known that for the VP ellipsis to be licensed, a strong parallelism must obtain between the antecedent VP and the elided VP.

For example the antecedent clause in (26) is ambiguous: the universal quantifier can take either wide or narrow scope with respect to the indefinite. However, the scope configuration which is selected in the first clause must also arise in the second clause (cf. Fox 2000 among others.).

(26) A doctor visited every patient and a nurse did too.

(25) GIANNI BEANS EAT-QUICKLY. PIERO SAME

However, the parallelism in (25) might be explained also under the VP anaphora analysis, if it is assumed that the antecedent VP structures the discourse, so the speed reading stops being contextually relevant if the amount reading has been selected in previous discourse (and the other way around).

All in all, the parallelism condition in (25) is consistent with the VP ellipsis analysis but does not discriminate between this analysis and the alternative analysis in terms of VP anaphora.

ANOTHER CASE OF PARALLELISM?

The parallelism requirement on VP ellipsis might also shed light on a fact that we do not fully understand yet. Our informants reject the continuation with SAME in sentences like (27), in which the verb EAT is used intransitively and the missing VP contains the verb but not the adverb modifying it. However the continuation in (28), in which the missing VP contains verb + adverb, is ruled-in. Also accepted is the continuation in (29), in which the missing VP contains verb + object, but not the adverb.

- (27) GIANNI EAT QUICKLY. *PIERO SAME SLOWLY
- (28) GIANNI EAT QUICKLY. PIERO SAME
- (29) GIANNI BEANS EAT QUICKLY. PIERO SAME SLOWLY

- (27) GIANNI EAT QUICKLY. *PIERO SAME SLOWLY
- (28) GIANNI EAT QUICKLY. PIERO SAME
- (29) GIANNI BEANS EAT QUICKLY. PIERO SAME SLOWLY

We speculate that the parallelism condition on VP ellipsis can only be met if the event that must be matched in the SAME clause is sufficiently specified. This happens if a manner adverb qualifies the eating event (cf. 28) or the Theme theta role is specified (cf. 29). In (27), this event specification does not take place and underspecification inhibits parallelism.

DUAL

We have been arguing that the SAME is an adverbial that can accompany VP ellipsis as in (30), rather than being a verbal anaphora as in (31) .

(30) Gianni bought a new computer and Mario did too

(31) Gianni bought a new computer and Mario did that too

However, there is no reason why LIS should not have a construction involving verbal anaphora. In fact, we think it does. The relevant construction involves the sign DUAL, which is often glossed in Italian as “duale”.

(32) is a discourse made by three independent sentences. The sign DUAL in the last sentence indicates that Maria ate both pizza and salad. DUAL refers back to the area of space where GIANNI and PIERO were articulated, so it is indirectly linked to the area where the VP “PIZZA EAT” and the VP “SALAD EAT” were articulated.

(32) GIANNI_i PIZZA EAT. PIERO_j SALAD EAT. MARIA DUAL_{i+j}
 “Gianni ate a pizza. Piero ate a salad. Maria did both things”

It is difficult to think that (32) is a case of VP ellipsis, since the coordinated VP that is allegedly elided (“ate a pizza and ate a salad”) is not present in either of the two sentences that precedes the sentence with DUAL.

(32) GIANNI_i PIZZA EAT. PIERO_j SALAD EAT. MARIA DUAL_{i+j}
“Gianni ate a pizza. Piero ate a salad. Maria did both things”

(32) GIANNI_i PIZZA EAT. PIERO_j SALAD EAT. MARIA DUAL_{i+j}
“Gianni ate a pizza. Piero ate a salad. Maria did both things”

It is more natural to think that DUAL is an anaphor that can take split antecedents, since anaphors can do that:

(33) I talked to Bill. Mary talked to John. Tony talked to both of them.

If DUAL is a verbal anaphor and SAME is an adverbial, they should be able to co-occur. In fact, they do:

(34) GIANNI IX-3_i PIZZA EAT. IX-3_j MARIA COCA-COLA DRINK IX-3_j. IX-3_k
 pause MAURO IX-3_k SAME DUAL_{i-j}

“Gianni ate a pizza. Maria drank a coke. Mauro got both a pizza and a coke”

CONCLUSIONS

We have identified a construction in LIS in which a VP can be omitted if there is a suitable linguistic antecedent.

Under standard assumptions, the linguistic antecedent counts as a constituent. Furthermore, we have shown that this constituent has a hierarchical organization.

So, sign languages, or at least sign languages like LIS, are not a counterexample to the claim that the hierarchical organization of the clause is a language universal.

Similarly, sign languages, or at least sign languages like LIS, are no special challenge to nativist theories of language that takes the hierarchical organization of the clause to be part of the biological endowment which allows a child to develop a language, spoken or signed.

Special thanks to...

Gabriele Caia
Rosella Ottolini
Mirko Pasquotto

REFERENCES

Fiengo, Robert and Robert May (1994) *Indices and Identity*. Cambridge, MA. MIT Press.

Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, MA: MIT Press.

Hankamer & Sag, 1976 “Deep and Surface Anaphora”, *Linguistic Inquiry*, 7:3,3 91-428