Development and Representation of Italian Light-*fare* Constructions

Valeria Quochi

Consiglio Nazionale delle Ricerche, Istituto di Linguistica Computazionale "A. Zampolli" Pisa, Italy

1   Introduction

Light, or Support, Verb Constructions (LVCs) have received extensive attention in the literature and have been studied from different theoretical backgrounds: generative grammar, relational grammar, lexico-semantics, lexical-collocational approaches (Grimshaw & Mester, 1988; Myiagawa, 1989; Namer, 1998; La Fauci & Mirto 2003; Alba-Salas, 2002; Giry-Schneider, 1987; D'Agostino & Elia, 1997; Cantarini, 2004; Sinclair, 1991; Moon, 1998; to mention just a few). Typical examples of LVCs are *give a talk, take a walk* or *tenere una conferenza* (lit. hold a conference 'make a speech'), *fare una passeggiata* (lit. do a walk). The mainstream position is to consider them as kinds of multi-word units, (semi-)fixed idioms or collocations; a view that reflects the bias of theories based on a neat separation between grammar and lexicon. Yet, they constitute a conspicuous class of constructs across even typologically different languages. Similar structures have been observed and investigated in

various world languages: English, French, Dutch, Japanese, and many others (see Alba-Salas, 2002 for a review of the languages investigated).

Although there is some variation depending on the theoretical approaches, the definition of LVCs is generally based on 3 main assumptions: 1) the subjects of the noun and of the verb/clause need to be co-referential, 2) the verb is semantically light or bleached, 3) the noun is the semantic nucleus of the sentence: it determines argument structure and assigns the semantic roles to the sentential arguments. By a strict definition usually adopted by generative approaches, only verbal nominalizations are allowed as the semantic heads of the constructs. A looser definition, common in relational grammar and lexico-syntactic approaches, requires the noun to be predicative, argument-taking, no matter its morphological status.

According to the stricter definition, thus, only expressions like *fare un salto* as in (1) would be true LVCs because *salto* 'jump.N' is a nominalization, while *fare una serenata* and *fare rumore* as in (2) and (3) would be LVCs only by the broader definition.

(1) Marco   fa   un salto

    Marco   does   a jump

    'Marco jumps'

(2) Romeo fa   una serenata   a Giulietta da   sotto   il balcone

    Romeo makes a   serenade   to Juliet   from under the balcony

    'Romeo serenades Juliet from under the balcony'

(3) Il   motore dell' aeroplano   fa   un rumore   assordante

    The engine of.the plane   makes a   noise   deafening

    'The plane engine makes a deafening noise'

In any case, LVCs constitute a challenge for syntactic theories, esp. those based on the centrality of verbs in determining syntactic and semantic structures, because it is not the verb that determines the number and kind of arguments at sentence level, but the noun. In many languages LVCs behave like lexical verbs at sentence level, but at the same time show a certain degree of syntactic freedom, semantic transparency and productivity, which makes it difficult to treat them purely at the lexical level. This is also the case in Italian.

Moreover, by both definitions, expressions like *fare le carte* 'to shuffle the cards' and *fare paura* 'to scare sb' would be considered as fixed idioms bearing no relationship to LVCs, because the noun is not (intrinsically) predicative (4), or the subjects are not co-referential (5) (cfr. Giry-Schneider 1987).

> (4) Ora   faccio   io   le   carte
>
>   Now   do   I   the   cards
>
>   'I am shuffling the cards now'

> (5) Mi   hai   fatto   tanta   paura
>
>   To-me   do.2SG   done   very much   fear
>
>   'You scared me very much'

Although it is true that (4) can be considered idiomatic and (5) presents some important (argument structure) differences from typical LVCs, they also share both structural and semantic similarities that get lost in these accounts.

Focusing mainly on structural or formal properties, the traditional approaches to LVCs, as we have briefly seen, show important limitations. In particular, they fail to account for the productivity and semantic relatedness of (some of) these constructs.

On the contrary, cognitive, usage-based theories of language, and in particular Construction Grammar (Goldberg, 1995 and 2006; Croft, 2001; Mirjam Fried & Ola Östman, 2004), provide an interesting

alternative for the investigation of these types of expressions especially for seeking an explanation of their productivity and relatedness. While not specifically addressing the issue of Light Verb Constructions, they seem to offer interesting tools to account for (local) regularities and to explain their partial productivity.[1] The notion of family, or network, of constructions (Goldberg, 1995, Croft 2001) inspired the present analysis as well as the attention given to acquisitional issues. Investigations of language development are, in fact, fundamental for cognitive-functional approaches to language, because their main and common assumption is that language is learned from the ambient language. Thus, if language is not innate and not hardwired in the brain, then we need to understand how children arrive at acquiring the linguistic structures and competence of adults, from their first rudimental attempts.

The ultimate goal of this contribution is thus twofold: first, I try to show that LVCs are better accounted for as a network of related constructions organized as a radial category; and second I intend to provide support to a functional, usage-based hypothesis of language acquisition and claim that there exists a LVC pivot schema that children learn at an early stage because it serves the purpose of naming new events, activities or situations for which they do not have a single lexical word yet.

To achieve these goals I make use of naturalistic data and study the distribution and development of constructs with the Italian light verb *fare* 'do' in Child Language and in Child Directed Speech. First, the dataset is thoroughly analyzed focusing on the properties of the nouns in the LVCs and then on the constructions as a whole. In particular, I will advocate that an account of LVCs in terms of family of constructions (Goldberg, 1995) is more appropriate and explicative than traditional approaches. So, a set of Constructions will be identified and coarsely defined based on the data. Finally, the study of the distribution and evolution in Child Language of the constructions identified will bring us to sketch a possible development pattern of the LVCs in child language, which is in line with functional

[1] To the best of my knowledge, very few studies within the framework of Construction Grammar specifically addressed support/light verb constructions (Palancar, 2003, Family, 2009, Doğruöz and Backus, 2009).

hypotheses of language acquisition and in particular with Tomasello's verb island hypothesis (1992, 2003).

2   A developmental study of *fare* LVCs using naturalistic data

The main focus of the research described here is to study the use of the light-*fare* + direct object constructs in early childhood and in Child Directed Speech (CDS hereafter) with the aim of exploring the nature of Light Verb Constructions broadly defined on a basis of spontaneous language production. A dataset of developmental data was semi-automatically created starting from the corpus of transcriptions of communicative interactions contained in the CHILDES database (MacWhinney 2000, and web site).

I have first approached the analysis of the data assuming the perspective of traditional views of LVCs, and thus expected that, being Light Verb Constructions (LVCs) close to lexical units (semi-fixed idioms in the traditional sense), they are stored, unrelated, rote-learned items both in adult and child language. This, in fact, is not the case, as it becomes clear from the observations made and the problematic cases found. The analysis of the data thus will lend support to a different explanation, i.e. to a functional-constructionist approach to language acquisition.

2.1   *The data*

The CHILDES database consists of a set of transcripts of "spontaneous" child-adult conversational interactions. Each transcription file represents one recording session. The Italian collection used in this study consists of longitudinal transcriptions of interactive sessions with eleven non-impaired Italian-speaking, children. For the present study, 4 corpora are used, which collectively cover an age span between 16 and 40 months and consist of a total of 117 transcriptions:

The Antelmi corpus is a longitudinal study of one girl observed from the age of 26 months to 40 months.

The Calambrone corpus contains both longitudinal and cross-sectional data from six normal and 11 language-disordered children. Only the data of the 6 non-impaired children, 4 girls and 2 boys, which cover an age span from 19 to 39 months, is used in this study.

The Roma corpus is a longitudinal study of one boy observed from the age of 16 months to 20 months.

The Tonelli corpus is a longitudinal study of three children, one girl and two boys, and covers an age-span from 17 to 26 months.

The transcriptions were semi-automatically analyzed and all utterances that contain a verb were stored in a database.

At this point, a first analysis of the distribution of light verbs and of utterance types is performed in order to assess the general properties and trends of both CDS and child language (CL). It turns out that CDS is richer in questions than in declaratives or imperative utterances and that *fare* is more frequently used in questions than in other utterance types. This is interesting as it might constitute a kind of priming for LVCs. Children, not surprisingly, use mostly declarative utterances. Notice that a similar trend has been observed in English by Cameron-Faulkner and colleagues (2003). After observing the general trends, for each session all CDS and CL utterances in which an instance of the verb *fare* 'do' co-occurs with a noun in direct object position were annotated and extracted. This constitutes the main data for the study discussed here. In the rest of the paper, we shall refer to this data as the *fare* dataset, or simply dataset. We shall also refer to "*fare + N* in apparent direct object position" as *fare* + noun combinations, constructs or expressions alternatively with a theory neutral attitude. The use of "construction" will always have to be intended as theory specific in Construction Grammar terms.[2]

---

[2] I.e. a Construction is a pairing of form and meaning

A simple statistical analysis of the dataset shows that, overall, *fare* + noun combinations produced by children account for 41% of all their uses of the verb *fare*, while for adults they account for 48% of their uses of the verb, as shown in   Table 1. [3]

Constructs	Child's production	Adults' production
Fare + noun	250 (41%)	1848 (48%)
All *fare* contexts	611	3855

Table 1: *fare* dataset composition

Looking more qualitatively at *fare* constructs in CDS, one striking albeit not surprising observation is that the verb *fare* very frequently has a generic semantics: it often occurs with no "heavy" argument or no argument at all, especially in questions. Examples are given in (6) below.

(6) a. Cosa fai ?

What do?

'What are you doing'?

b. Come fai a?

How do.2SG to?

'How do you …?'

c. Fai   così

do.2SG   so

'Do it like this'

---

[3] In a preliminary analysis of the general characteristics of the dataset it was observed that there is no consistent qualitative difference between verb production in mothers and other adults, and therefore both have been used as representative of Child Directed Speech.

Qualitatively, it is also worthy of note that *fare* in such contexts seems to elicit descriptions or mentions of actions, events or situations in response, and that children usually start answering by repeating the verb itself (which is also phonologically salient in the speech stream) and add other words that evoke the situation or action prompted by the adult, although not necessarily in a syntactically adult-like utterance.

3   *Fare* constructs. Three macro categories

As mentioned above, most studies on LVCs are based on three main assumptions: 1) the subjects are coreferential, 2) the verb is semantically light or bleached, 3) the noun determines argument structure. As such, traditional approaches require that the noun be predicative: i.e derived from a verb, or argument taking (which admits some abstract nouns).

Assuming these assumptions as correct, I started by analyzing our dataset on the basis of the types of nouns occurring in direct object position of the verb *fare*. In this exercise, thus, the utterances are annotated according to the supposedly "inherent" formal (lexical semantic or morphological) properties of the nouns, as if out of context.[4]

I thus distinguished three basic types of nouns: nouns morphologically related to verbs, abstract (thus potentially predicative) nouns, and concrete nouns (Table 2 below gives their distribution in the dataset). Now, if we think in terms of the traditional approaches to LVCs the expectation is to find (most) LVCs in the group of constructs with nouns morphologically related to verbs, some LVCs among the constructs with abstract nouns and no LVCs among constructs with concrete nouns. As will become clear in the following subsections, this is not the case. Although certainly few in number, there are expressions in the concrete noun group that bear no literal meaning but rather interesting

---

[4] This is clearly controversial. In fact, the objections and problems arising from this analysis contribute to support a constructionist approach to LVCs.

similarities to classical LVCs. As such they would be treated as isolated idiomatic expressions in traditional, formal approaches, while in construction grammar their similarity to a wider class of expressions could be accounted for.

	Concrete nouns	Verb-related nouns	Abstract Nouns	All nouns
Adults	1006 (54%)	302 (16%)	540 (30%)	1848
Children	146 (58%)	35 (14%)	69 (28%)	250

Table 2**:** Noun Types in the *fare* dataset (Token Frequency)

But let us proceed step by step and analyze the constructs by noun groups.

3.1   *Fare + Nouns Morphologically Related to Verbs*

This class groups together nouns that are morphologically (or perceptually) related to common verbs, no matter the direction of derivation since it is not easy to determine the direction of the derivation, especially in the case of zero derivation, and in any case it seems to be more a matter for linguistic description than for the speaker's consciousness (cfr. Iacobini 2000).

For this task I had to make some other decisions as to what to classify as verb-related nouns, which may not be uncontroversial. So, for example, nouns like *fotografia* 'picture' are included in this group because the corresponding verb *fotografare* 'to take a picture' is quite commonly used in everyday language, while *triangolo* 'triangle' is not included although it corresponds to the verb *triangolare* 'to triangulate', because the verb is a technical term.

Within this group I included all types of derivation: zero derived nouns (further separating forms altered by diminutive or augmentative suffixes), *-ata* nominals, *-ione* nominals, and nouns derived by means of other suffixes. The most interesting quantitatively are the first two.

Zero Derived Nouns are derived by conversion directly from one form of the base verb (e.g. salto 'jump', gioco 'game', *ballo* 'dance') and are considered to be highly conservative of the properties of the base verb. However, they are for the most part countable and therefore are cognitively closer to nouns than verbs (cfr. Gaeta 2002). If we take their semantics into account, they generally denote some type of action, and thus belong to the traditional class of Nomina Actionis, or to the Process/ Result alternation classes.

-ata Nominals are formed from the past participle of the base verb and always in in feminine form (e.g. *passeggiata* 'walk', *girata* 'stroll'). The *-ata* derivation from regular verbs is highly productive and transparent. Semantically, they denote single occurrences of the action or event denoted by the base verb; they cannot normally be used in a generic sense, but only as individuated instances, and are generally not compatible with definite determiners, except in marked contexts where the deverbal is closer to an entity-denoting noun. Their characteristics are highly compatible with "classic" LVCs. In general and in line with our expectations, the constructs in this group appear to be cases of canonical LVCs: e.g. *fare un salto* do a jump', *fare la spesa* 'go shopping' as in examples (7) and (8). As is conventional in the CHILDES database, in the transcripts CHI always refer to the child, MOT to the mother and ADU to other adults.

(7) CHI: oa   io faccio i   satti   dice   Ila.

        Now   I   do   the   jumps   says   Ila

        'Now I jump, says Ilaria'

                       [Rosa, 2:11]

(8) MOT: questa è la   nonna che   va   a fare la   spesa.

    This   is   the grandma who goes to do   the   expenditure

    'This is the grandma who goes shopping'

             [Rosa, 2:5]

These can be considered classic examples of LVCs with the noun contributing most of the semantics to the clause and the verb being semantically "light". The schematic abstract semantics of such constructs appears to be something like DO ACTION, provisionally labelled: 'Perform Action'.

In Table 2 above we saw that *fare* constructs containing nouns morphologically related to verbs account for the 16% of all adult *fare* + nouns constructs, and 14% of children's, that is: the production of *fare* + verb-related nouns in children is proportionally very similar to adults', which suggests that children do not creatively produce such expressions, but rather imitate what they hear in their environment. This is in fact corroborated by further inspection of the data.

Table 23 below reports the frequency distribution of the two main derivation types considered.

	Adult			Child		
	Types	tokens	Ratio	types	Tokens	Ratio
Zero-derived nouns	33	159	*0.21*	13	19	*0.68*
ata-nominals	17	60	0.28	3	8	0.38
Other deverbals	21	83	0.25	3	8	0.38
total	71	302	*0.24*	19	35	*0.54*

Table 3: Distribution of various subtypes of verb-related nouns in the dataset. Both type and token frequency is given as well as TTR

As we see, children mostly use zero derived nouns, while adults pool from a wider range of types. Still adults seem to show little variability, with a low type/token ratio that indicates the repeated use of relatively few lexicalized constructs, which is confirmed by qualitative analysis.

Overall, however, we observe a similar trend both in CDS and in child language in terms of token production: children's production of zero derived nouns account for 7% of their *fare* +noun combinations and for 54% of their combinations with a verb-related noun, against the 9% and 53% in adults.

Children produce fewer types and relatively more tokens than adults of constructs with –*ata* , e.g (9), and nouns derived by means of other suffixes.[5] Again, this is not surprising since they are morphologically more complex and thus expected to be learnt at later stages.

> (9) CHI: perche' si va a fare una girata, allora devono stare pronti
>
>   because we go to do a turned, then must.3PL stay ready
>
>   'because we are going to take a stroll, so they must be ready'
>
>   [Camilla, 3:01]

The relatively low type-token ratio (TTS) with nouns derived by suffixation also suggests that these productions are rote-learned, which is in fact supported by an inspection of the data: in all cases the constructs used are conventional and entrenched.

Still, even though they are conventional or lexicalized nominalizations, these nouns always keep a process meaning similar to their base verb, and this meaning may be what licenses the LVCs in the first place. For example, *Spesa* lit. spent, expenditure 'shopping' is lexicalized, but may still refer to an activity which is metonymically related to the meaning of *spendere* 'to spend'.

---

[5] The constructs are: *fare+passeggiata* 'do+walk', *fare+girata* 'do+stroll', *fare+spesa* 'do+shopping', *fare+carezza* 'do+caress', *fare+foto* 'do+picture'.

In order to get a clearer idea of the evolution of the use of such constructs, we now look at the distribution of *fare* constructs with verb-related nouns in children across age groups, in Figure 1: ***Fare* + verb related nouns across age**. In the figures of the rest of the paper you will find, on the x axis, the age of children expressed in months; on the y axis the frequency. The dark line represents CL, the light one CDS. The same graphic applies to the dotted lines that indicate the linear trends.
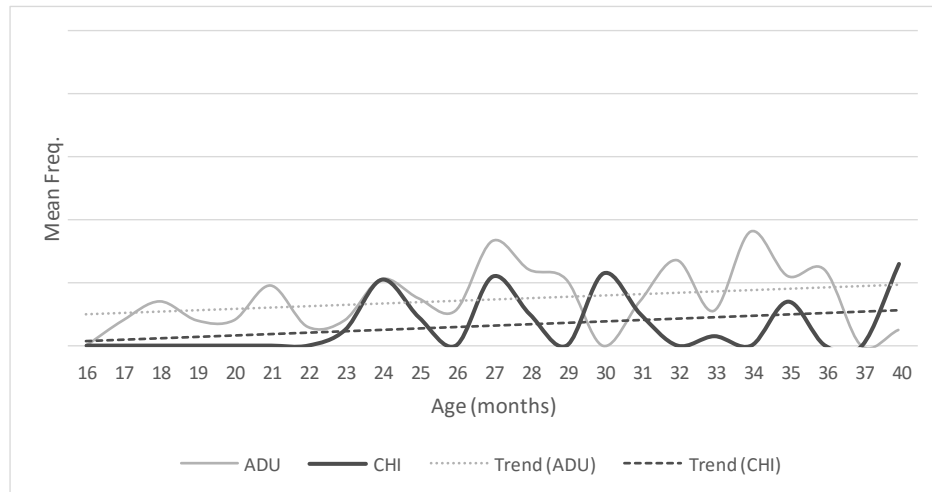

Figure 1: *Fare* + verb related nouns across age

Here, children start producing *fare*+V-related noun constructs around 24 months and tend to increase their production in the following periods, though not dramatically. However, such a growth in production is likely dependent on the general vocabulary growth which takes place in the time span covered by this research and on the fundamental role of the input on the child's language, as predicted by usage-based, functional theories of language acquisition. The distribution of the same constructs in CDS shows a similar tendency. Vocabulary growth, moreover, is generally considered an important factor that encourages a process of abstraction into more general classes (see i.e. Bates & Goodman 1997; Ninio, 1999).

Around 30-31 months of age we observe a counter-tendency: while adults do not produce virtually any combination with verb-related nouns, children are productive. The constructs uttered by children

in those sessions contain the nouns *bagno, foto, carezza, disegno* 'bath, picture, caress, drawing'. As observed above, however, these are lexicalized nominalizations, the constructs are rather conventional and all types observed are also used by adults in previous sessions. Also, in most cases it is the same child that produces the same construct type with a relatively high token frequency. All these observations favor the assumption that *fare* constructs with verb-related nouns are rote-learned and meets the expectation to see imitative, not creative child production.

3.2 *Fare + (Physical) Entity-denoting nouns*

This group of nouns co-occurring with *fare* contains typical (physical) entity-denoting nouns like *casa* 'house/home', *balena* 'whale', *sedia* 'chair'. As such, they are found mostly in canonical transitive syntax, in which the verb is used as a fully lexical verb of creation. Consider for example (10) where the verb can be paraphrased with *disegnare* 'draw' as in (10a.), or *costruire* 'build/make', *decorare* 'decorate' as in (10b.).

(10)   a. CHI: faccio   una balena grossa!

           Do.1SG a   whale big   !

           'I'm making/drawing a big whale'

                                [Raffaello 2:7]

        b. MOT: chi   l' ha fatto l' albero di Natale?

             Who it has done the tree   of Christmas ?

             'Who made/decorated the Christmas tree?'

                                [Marco 2:0]

Looking at the distribution of *fare* + Concrete Nouns across age groups for both children and adults (Figure 2), in spite of the great variability, there is a constant growth of these constructs in CL (as indicated by the trend line), whereas adults show a more constant behavior.
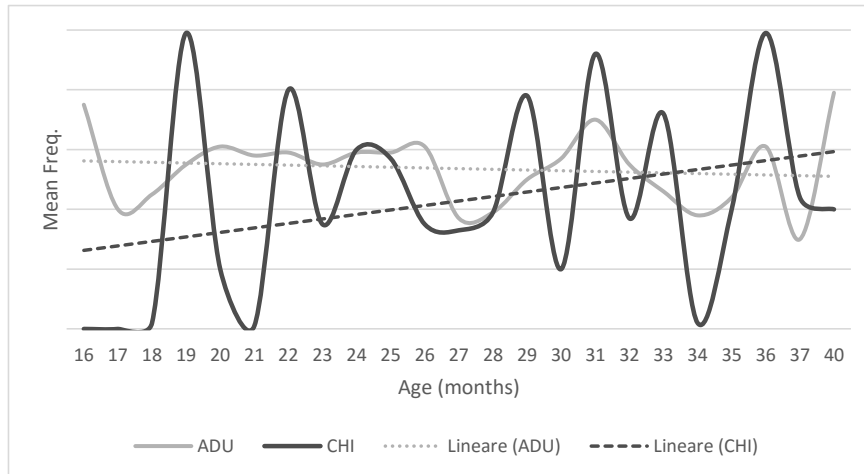

Figure 2: *Fare* + (Physical) Entity-denoting nouns across age

Children, not surprisingly, prove to be productive in this group of constructs and the trend observed is in line with our knowledge of the development of the lexicon and syntax in early childhood: productive transitive constructs appear after the one-word utterance phase, and after the first noun-noun and noun-verb combinations (Camaioni 2001, Tomasello 2003, Lust 2006).

In the literature on LVCs, constructs with such nouns are generally excluded *apriori* from the analyses. In this study I decided to take them into consideration, given the possibility that, in combination with *fare,* they give rise to lexically filled LVCs, and therefore instances, albeit non-prototypical, of the general schema. If we consider for example (11), although the noun 'little horse' is intrinsically clearly entity-denoting (it is even a diminutive form), the construct does not refer to a creation event in a physical sense: there is no entity coming into being, rather the whole construct

refs to an activity (i.e. the game of taking a child on ones lap and moving ones legs *as if* to be a horse).

> (11) CHI: Luca   si   fa   valluccio ?
>
>   Luca   SI   do.3SG   horse.DIM
>
>   'Shall we play the little horse, Luca?'
>
>   [Diana, 1:11]

While this is clearly a lexicalized expression, the strategy for evoking an activity or situation is strikingly similar to that of canonical LVCs, that is with verb-related nouns. In these cases, the semantics of the verb *fare* seems closer to PERFORM than to CREATE/MAKE.

Although most of the constructs in this group involve some sense of creation, where the verb can be paraphrased with verbs like 'drawing', 'building', 'decorating', the fact that there are some metaphorical, or idiomatic, expressions sharing strong similarities with canonical LVCs provides us with interesting support against traditional accounts of LVCs.

3.3 *Fare + Abstract Nouns*

Many of the verb-noun combinations in this group are expected to be LVCs according to the broad definition, since many abstract nouns can be considered/are observed to be argument-taking.

Typical examples of *fare* + abstract nouns found in the dataset are: *far bang*, *fare acrobazie*, *fare chiasso* (lit. 'make/do bang, acrobatics, noise').

Most of the combinations in this group yield expressions bearing some kind of non-literal meaning: some of them are clear examples of LVCs, in which the verb does not contribute the usual creation meaning, but a more abstract, general one (ex. *fare rumore, ginnastica, paura, blitz* 'make/do noise, gymnastics, fear, blitz'). Some are typical "motherese" expressions that are not normally used in adult conversations, but are nevertheless conventional (ex. *fare totò* 'to spank', whose noun is an entry in

the De Mauro dictionary). Others are considered as (semi-)fixed idioms like *farsi una cultura* lit. make oneself a culture 'become educated', *fare (i) guai* lit. make (the) mess 'to mess up', or *fare il proprio repertorio* lit. Make the own repertoire 'to do/perform one's own repertoire' as in (12).

(12) ADU: Giulio fa   tutto il   suo repertorio

Giulio makes all   the his repertoire

'Giulio makes his whole repertoire'

[Viola, 2:1]

In this macro-group, two classes of expressions deserve special attention: those related to *sound emission* (e.g. *fare un verso, fare rumore, fare chiasso* 'make a cry, a noise', to moo'), which are LVCs in the broad sense; and those related to *gestures* (like *fare ciao, fare caro, fare totò* (lit. make hello 'say hello', do dear 'caress', do slap 'to slap'), which do not fall within the definitions of LVCs, but share formal and semantic properties with them and seem to play a role in the learning of more typical constructions. Among the similarities with expressions that, instead, meet the definitions of LVCs perfectly, I mention here that these expressions can be paraphrased using a single verb lexeme, although the nouns are neither 'intrinsically' related to such verbs, nor intrinsically argument-taking. *Caro* 'precious, dear', for example, is more commonly used as an adjective and it's relation to the noun *carezza* 'caress' and the verb *(ac)carezzare* 'to caress' is more an etymological issue than an everyday speaker awareness.

Regarding the distribution of constructs with abstract nouns, it is surprising to find high type frequency, like that with concrete nouns, among constructs with abstract types as well, since children are usually thought to acquire concrete words first, and abstract ones later. We may speculate that this is an effect of their use in the specific context *fare*+Noun and that their distribution in the overall

corpus would confirm the shared knowledge about lexicon acquisition. If so, abstract nouns must play a special role in these constructs.

As for the development of such constructs, a quite different behaviour than with verb-related nouns emerges. There is a mild growing trend in children here (see Figure 3), but also a greater variability as shown by the higher type frequency. Children start producing these constructs around 20 months of age and they seem to acquire some creativity around 24-25 months. Their frequency is higher than the frequency of the constructs with verb-related nouns, and higher than in CDS.
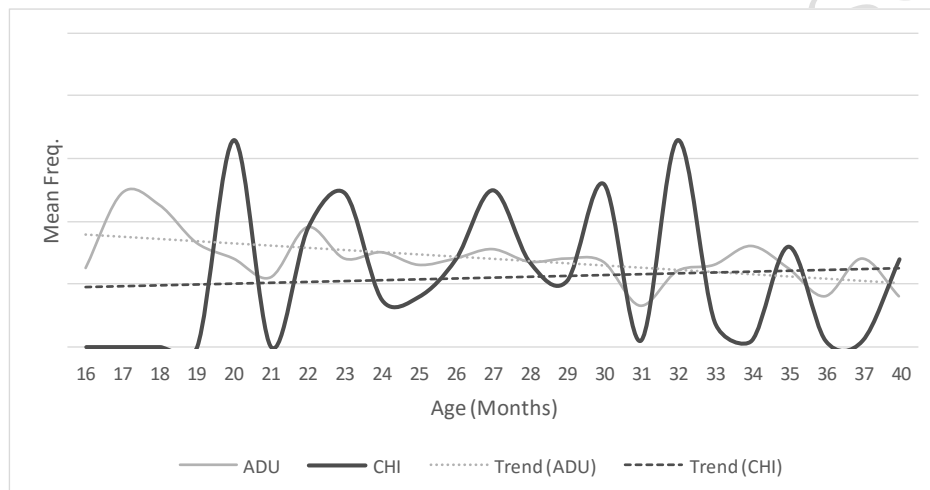
Figure 3: *Fare* + Abstract nouns across age

What is also interesting here is that children produce more constructs with abstract nouns than adults from the age of 20-24 months up to 33 months. After this period, their production decreases again. Turning to a qualitative inspection of the types of noun in such constructs, we find conventional "baby-talk" expressions like *fare la nanna* 'take a nap' or *fare paura* 'scare', but also "invented" expressions with onomatopoeic words like (13) and (14).

(13) CHI: sopra una collina # abbaiando # fanno   bu bu .

On   a   hill   barking   do.3PL   bu bu

'on the top of a hill # barking # they do "bow wow"

[Raf, 2:8]

(14) CHI: no guarda uno uno sparato e   loro fanno pum !

No look   one one shot   and they make pum

'No, look one is shot and they shot him'

[Raf, 2:11]

Some of them turn out to be the child's "inventions" that do not survive in adult language, and are not even reflected in the input.

These inventions are very interesting because they share surface properties with canonical LVCs, the verb does not seem to be used in its basic lexical sense and the entire construct often evokes whole situations or denote actions and activities. For the moment, we shall call this class of expressions the *fare* + SOUND pattern.

An example of child 'misuse' or invention is *fare musica* 'make music' (15), which is used by the child to mean that some object like a radio is emitting music, and not that a volitional agent is playing an instrument, as it might be used in adult language.

(15) CHI: ma   non   fa   musica?

But   not does   music ?

'Doesn't it [the radio] make music?'

[CAM, 2:6]

3.3.1   Onomatopoeic nouns as separate class

Since during the qualitative analysis of constructs with abstract nouns, onomatopoeic words (i.e. words indicating typical sounds of objects or cries of animals, words whose sound/phonetic iconically resembles the sound denoted) emerged as a conspicuous class, they have been further isolated as a subtype of the abstract noun category.

All onomatopoeic words used as surface direct objects of *fare* in the present analysis will be considered as nouns for reasons of consistency, given that many of the onomatopoeic words found in the corpus are also defined as nouns in traditional dictionaries (e.g. the De Mauro dictionary of contemporary Italian for example has a nominal entry for *miao* 'meow', the cry of cats).

Some of these words are sometimes used metonymically, especially in CDS, to refer to concrete entities, e.g *coccodè* 'cluck-cluck' for hen, *ciccì* for bird.

Children, not surprisingly, seem to prefer these words and show early productivity. *Fare* combinations with onomatopoeic words in CDS overall account for 8% of the abstract noun group, while in children it accounts for 13%.

In our dataset, children produce various types of these combinations from quite an early age. Results from a session per session analysis show that starting from 17 months, all children produce some of these construct types without directly imitating the adults. Also interesting is the fact that, while adults use more conventional onomatopoeic words (e.g. *coccodè, chicchirichì, dindon* 'cackle, cock-a-doodle-doo, dingdong'), children are more inclined to invent them (*bumbe* 'boom', *cià cià* 'splash', *crac denk …*).

Sound words in combination with *fare* are also often used to express events or activities, especially by children: Marco at 20 months says *fare ahm* meaning 'to eat', and Marco's mother, when he is 17 months of age, says *fare crack* meaning 'to break' in its two possible meanings, both as a causative verb (16) and as an inchoative (17).

(16) MOT: puoi   fare crac con queste, puoi   romper-le .

      Can.2SG do crack with these, can.2SG break-them

      'you can make crack with these, you can break them'

[Mar09]

(17) MOT: la scatola di polistirolo e' fragile e   ha fatto crac .

      the box   of polystyrene is fragile and has done crack

    'the polystyrene box is delicate and it broke'

[Mar09]

One meaning of *fare crack* implies that an agent intentionally breaks a patient object, while in the other meaning, it is the patient that *fa crack* 'breaks'. *Crack*, therefore, is metonymically related to the event implied, in that it is the "typical" sound of a breaking thing, and its combination with the verb *fare* in an iconic fashion suggests that this may be the way a child first learns LVCs.

All this seems to suggest that *fare* + N acts as a pivot for naming many "sound" and "action" events, especially in CL.[6] If we also consider that *fare* is very frequent in adults' questions, the role of this verb as a pivot for constructions naming actions and events may also be explained with a priming effect. Adults, especially mothers, appear to use *fare* in questions that aim at eliciting answers from their children about what activity or action he/she, or somebody else, is doing, or about what is happening. So, it is reasonable to assume that *fare* becomes a very general "action" word in early CL: a word that can combine with virtually any other word that activates or refers to a whole event or scene.

---

[6] We have encountered at least one example of CDS in which the *fare* + sound pattern is used to introduce a probably new lexical verb to the child, e.g. *fare crac* ' do crack' for *rompere* 'break' in (23).

4   A constructionist hypothesis for *fare*+N constructs

On the basis of the observations reported in the previous sections and the properties of the nouns and constructs in the various groups, it is now clear that a traditional formal approach to LVCs does not account for similarities among expressions whereas a constructionist explanation appears more viable. In particular, I shall claim here that the *fare*-constructs found in our dataset are better accounted for as a family of constructions. [7]

Taking into account the distribution of the constructs in CDS and some semantic, denotational properties of the nouns and constructs as a whole, I provide here a tentative representation of both the schematic constructions and of the family as a whole. It is not the goal of the present work to formally establish a full specification of the constructions' syntactic, semantic and pragmatic properties as well as of the specific relations among them. This is certainly an important and interesting issue for a constructional description of Italian, but is reserved for future research. The purpose of this study is, instead, to provide a tentative coarse-grained representation of the family of *fare* constructions as it emerges from the dataset, in an attempt to prove that such a representation is not only viable, but also that it can account for both the more schematic and productive types of LVCs and for the more conventional, lexicalized instances in a uniform way. I proceed, therefore, to give an informal description of these constructions, based on the exemplars found in the dataset.

Four more general, or schematic, constructions were identified that can account for several exemplars in the dataset, and tentatively assume, based on the distributional evidence discussed in section 3, that there is one central construction. Figure 4 displays a naïve representation of the *fare*-family of constructions.

The seemingly central construction, which we shall call the *Perform Intransitive Action* construction, would have a form-meaning like in (18):

---

[7] For the notion of family of constructions see in particular Goldberg (1995) and (2006).

(18)

	agt	<Activity/action>	(loc/goal)
V	Subj		(Comp)
Fare	NP	predN	PP

Pragmatically, the construction embodies the "EVENTS ARE OBJECTS" metaphor, which seem to be inherited by most of the related constructions (for details on the metaphor see Lakoff, 1980).

This construction is still quite abstract, and we may envisage that many single conventional LVC are individually stored as instances of the more general pattern, or of one of its subtypes, in the speaker's mind. Thus, *fare una passeggiata*, *fare una girata*, *fare un salto* 'take a walk, take a stroll, do a jump' would be individually stored. In addition to the higher token frequency of conventional instances of the LVC, a motivation for their storage might be that in most cases they are associated with a lexical verb expressing approximately the same content, and this synonymy link is likely to be explicitly represented[8].

---

[8]There is another possibility though: in Frame Semantics terms, an action noun and the corresponding verb, if any, would be part of the same frame; their semantic relatedness, therefore, could be already established at the lexical construction level. This would save storing many instances. Unfortunately, it is outside the scope of this work to enter into the details of Frame Semantics, or of how this could be integrated in a constructional representation of LVCs.
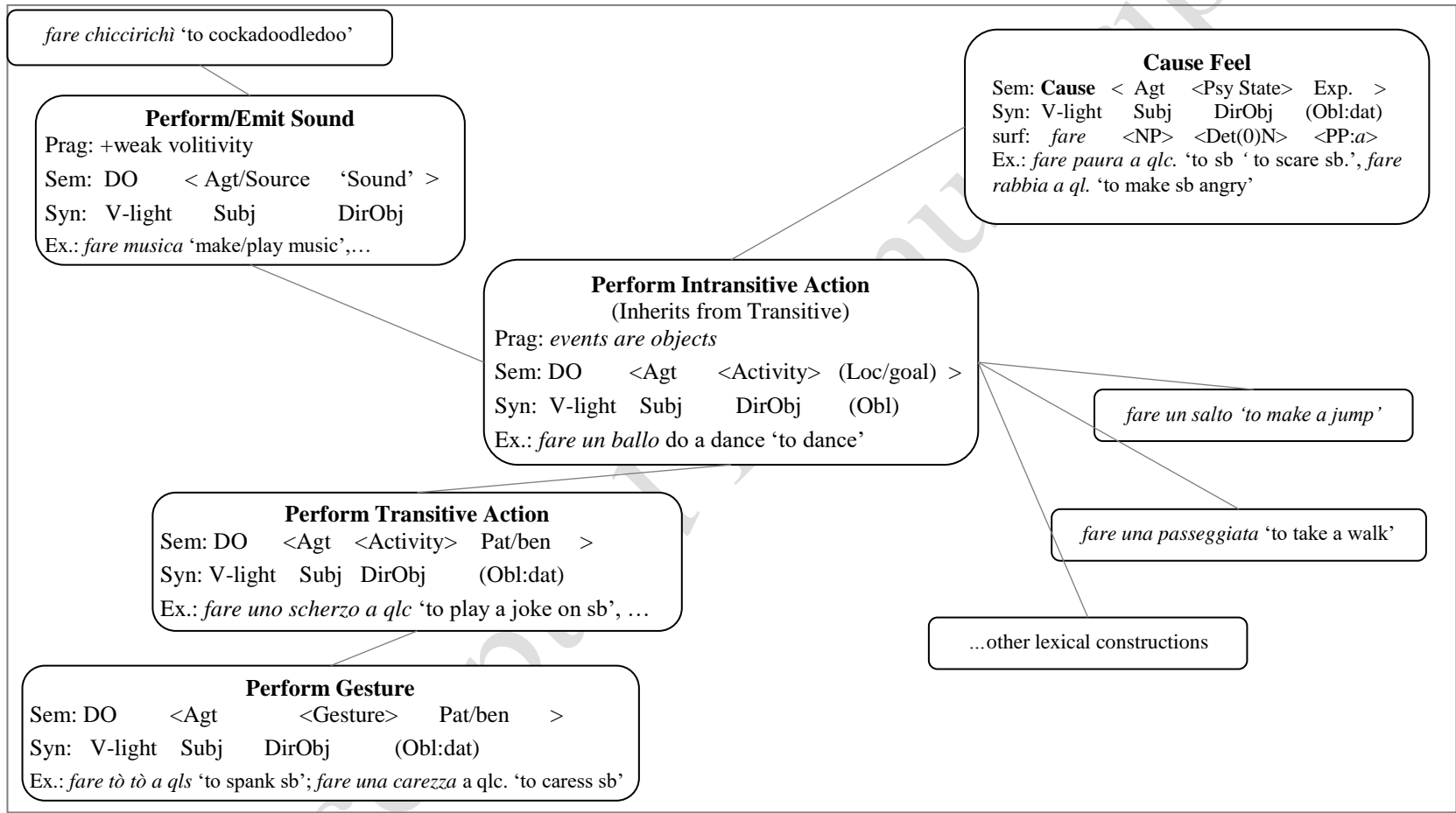
Figure 4: The *fare* LVC Family of Constructions

The *Perform/Emit Sound* construction roughly corresponds to the *fare* + Sound constructs discussed in section 2: e.g. *fare chiasso, fare miao* 'make noise, to meow'. This construction can be roughly described as in (19).

(19)

	Agent/source	<Emit/produce Sound>
V	Subj	
Fare	NP	predN

Examples of this constructions are *fare un fischio*, *fare bee* (Lit. do a whistle, do baa).

The *Perform Transitive Action* construction corresponds to constructs with nouns denoting a two argument, 'act upon' action. For example: *fare un colpo*, *fare una carezza*, *fare uno scherzo (a qualcuno)* 'hit, caress, spank, trick sb' (Lit. to do a hit, do a caress, do a joke to sb). The form and function of this construction would roughly be as in (20).

(20)

	agt	<Activity>	pat/ben
V	Subj	DObj	Obl
Fare	NP	NP	PP:a

The *Cause Emotion* construction corresponds to constructs of the type *fare* +noun denoting a psychological state (e.g. *fare paura a ql.* Lit. do fear to sb, ' to scare sb.'; *fare rabbia* Lit. do anger to sb, 'make sb angry) and its form and meaning might be something like (21).

(21)

	Cause	<PsyState>	Exp
V	Subj		Obl

| *Fare* | NP | Det(0)PredN | PP:a |

Even with these sketchy definitions of the properties of the constructions, the similarities among them appear evident. Treating *fare* constructs as a family of constructions, or a network of nodes organized like a radial category, allows us to capture the inheritance and similarity relation among them, as well as their differences, while maintaining the possibility of a full representation of conventional, entrenched instances, which can be stored as "peripheral" items. At the same time it allows us to represent lexicalized conventional instances as individual stored items.

Now, if we consider the predicative function of the noun as a dynamic property emerging in context, which can be explained and motivated on the bases of similarity and analogy, constructs with (apparently) concrete nouns are not excluded *apriori* from the *fare* LVC family, since they can still be forced into a predicative reading by the construction itself. This permits, for instance, relating expressions like *fare il trenino* 'do a conga line' with the LVCs "event naming" function.

5   Acquisition of the family of LVCs

After sketching a representation of the *fare* LVC family of construction as emerging from CDS, we now focus on the distribution of the constructions in CL, in order to understand how children acquire such constructions and whether they use them productively or not. The focus will be on the three most interesting constructions: the Perform Intransitive Action, the Perform Sound, and the Perform Transitive Action, and look at their distribution in the dataset and their evolution in child language. To this end, each *fare* + Noun occurrence in the dataset is further annotated as an instance of one of these three constructions, or none.

The three subsets of constructs obtained with this annotation partly overlap with the sets based on the types of nouns discussed above, but are not identical. Here, the main criterion is constructional form-meaning, as broadly defined above; therefore, in the same construction we may group abstract, verb-

related and concrete nouns together. *Fare canestro* 'score a basket', for instance, is considered and counted as an instance of the Perform Intransitive Action.

5.1   *The Perform Intransitive Action Construction*

Instances of Perform Intransitive Action are, for example: *fare il bagno* 'have a bath', *fare una scalata* 'make a climb', *fare il solletico* 'to tickle', *fare finta* 'pretend', *fare confusione* 'make noise/a mess'. This construction appears to be particularly frequent in both CL and CDS: overall it accounts for about 30% of all *fare* + Noun constructs in CL, and about 27% in CDS. If we look at the distribution in Figure 5, after 24 months we observe basically the same trend in children and adults: child production is clearly growing, but it basically reflects adults' uses.
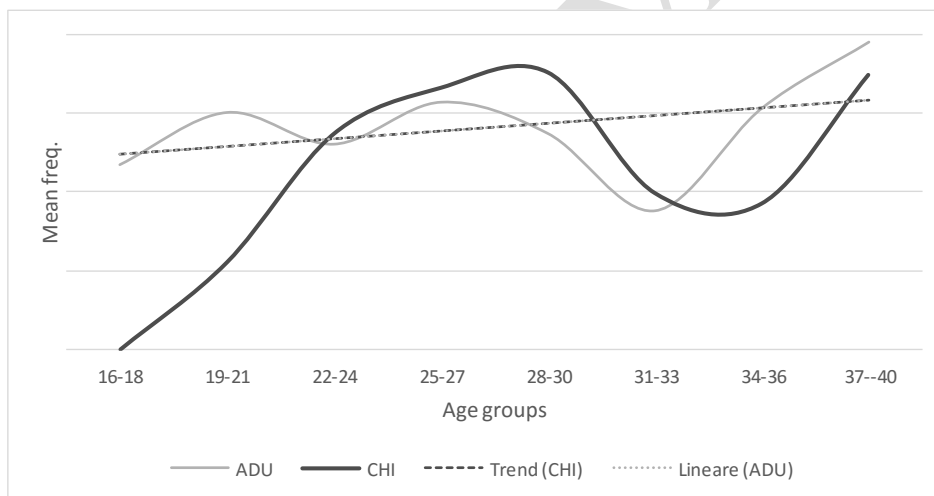
Figure 5: Development of the Perform Intransitive Action construction

This result is strikingly similar to the one found for verb-related nouns in section 3.1 above. Given that verb-related nouns constitute a considerable subgroup of this construction, it is not very surprising. If we exclude instances with verb-related nouns from the present set, we obtain the

distribution in Figure 6, which shows a quite different situation. Instances of this construction with abstract nouns are: *fare canestro, fare patatrak* 'score a basket, make a damage' as in (22) and (23).

(22) CHI: fatto canetro ho   fatto canetro.

        done basket have.1SG done basket

          'I scored a basket, I scored a basket'

                              [Mar26, 2:1]

(23) CHI: ho   fatto una patatrack io .

        Have.1SG   done   a   patatrak   I

        'I've made a mess'

                             [Mar24, 2:0]
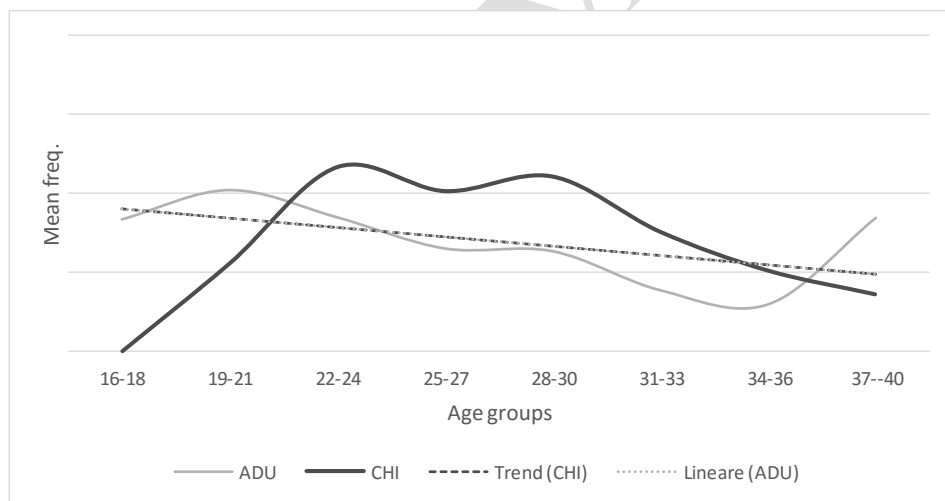
Figure 6: Development of the Perform Intransitive Action construction with abstract nouns only

Child production starts quite early, but until around 23-24 months children produce only highly conventional expressions like *fare (la) nanna* 'go to bye-bye'. At 24, instead they start to produce

proportionally more instances than adults, which shows that the pattern has been, or is being, acquired. Interestingly, children's production of the Perform Intransitive Action construction with abstract nouns progressively decreases after 29-30 month of age, as if they had experimented enough with it. Moreover, around 34-36 months of age, when the proportion of constructs with abstract nouns is quite low, we register an increase in the production of instances of the same construction, but with verb-related nouns (see Figure 7). This latter observation is again in line with common knowledge of language acquisition. Verb –related nouns, being more complex, are expected to be acquired later than other nouns. It is however interesting to observe a sort of mutually exclusive usage pattern between abstract and verb-related nouns.
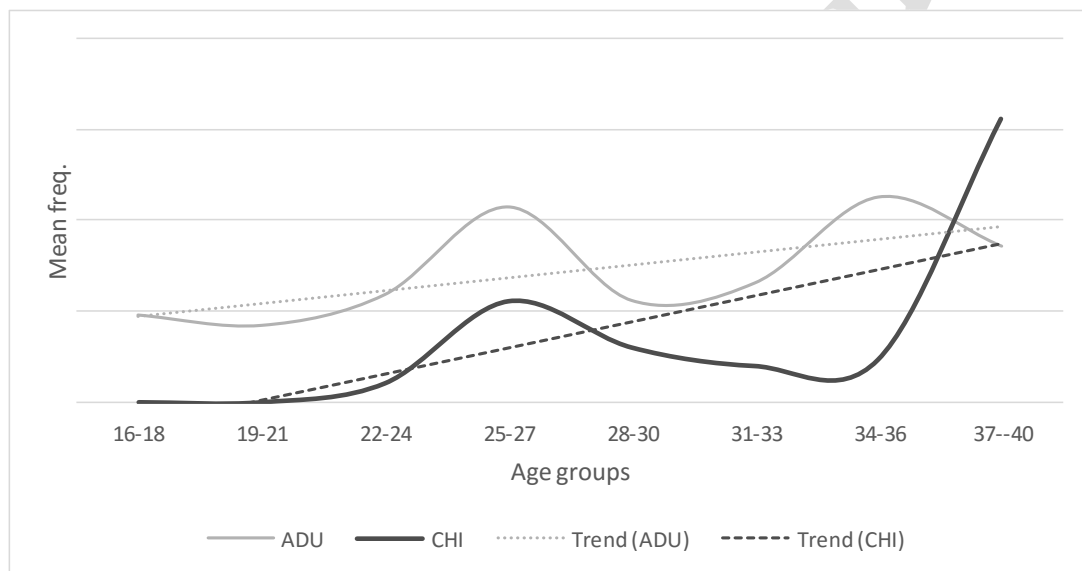
Figure 7: Development of the Perform Intransitive Action construction (verb-related nouns).

5.2 *The Perform Sound Construction*

The *Perform Sound* construction is a more specific construction that inherits most of its properties from the central one (the Perform Intransitive Action construction). Its distribution overlaps mostly with the class of constructs that we called Perform Sound in section 3.3. Typical examples of this

construction are: *fare* + onomatopoeic noun combinations like *fare muh*, *fare pum* 'to Moo, to shoot';
and other combinations with abstract nouns like fare *musica*, fare *rumore* 'make music, make noise'.
As we can see in Figure 8, children start to produce the *Perform Sound* construction quite early in
development, around 21 months and seem to be quite productive from around 23 months of age.
Around 33 months we see a significant decrease in production. Interestingly, adults show the opposite
tendency: they produce a high percentage of instances when their children are very young, and
progressively decrease their number as they grow. This trend in part explains why children start
producing these types of constructs earlier than others and the fact that they seem to apply the *fare*
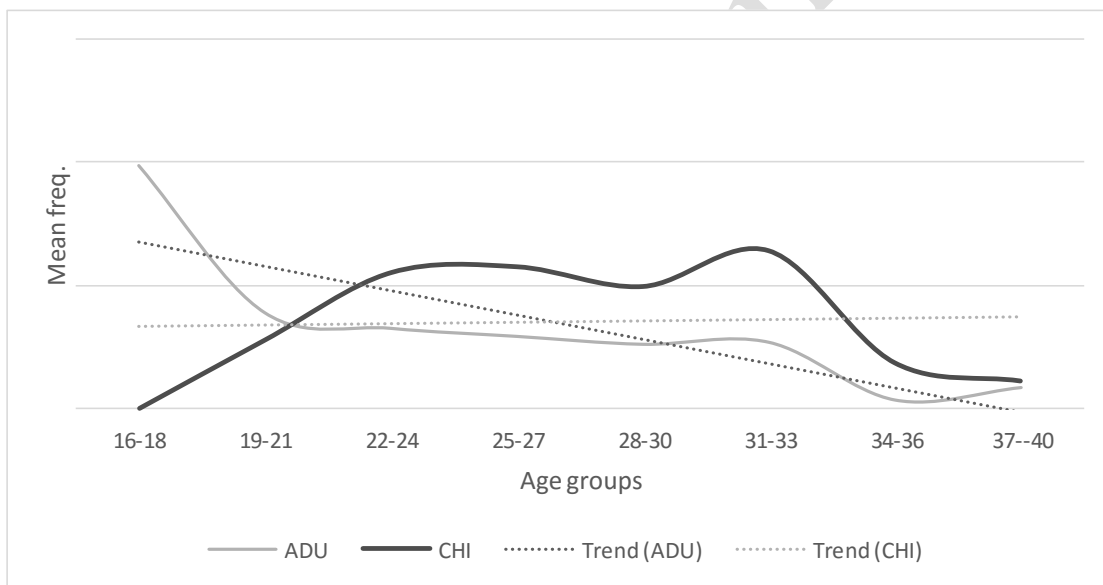ACTION generalization by using this type, as will be discussed later.

Figure 8: Development of the Perform Sound construction

5.3   *The Perform Transitive Action Construction*

Typical instances of this construction are: *fare la bua, fare caro/a, fare il solletico, fare paura* 'to hurt sb., to caress, to tickle, to scare'. In Figure 9 we see that while adults' production is quite constant, children's is rather discontinuous, which does not permit identification of a clear trend.
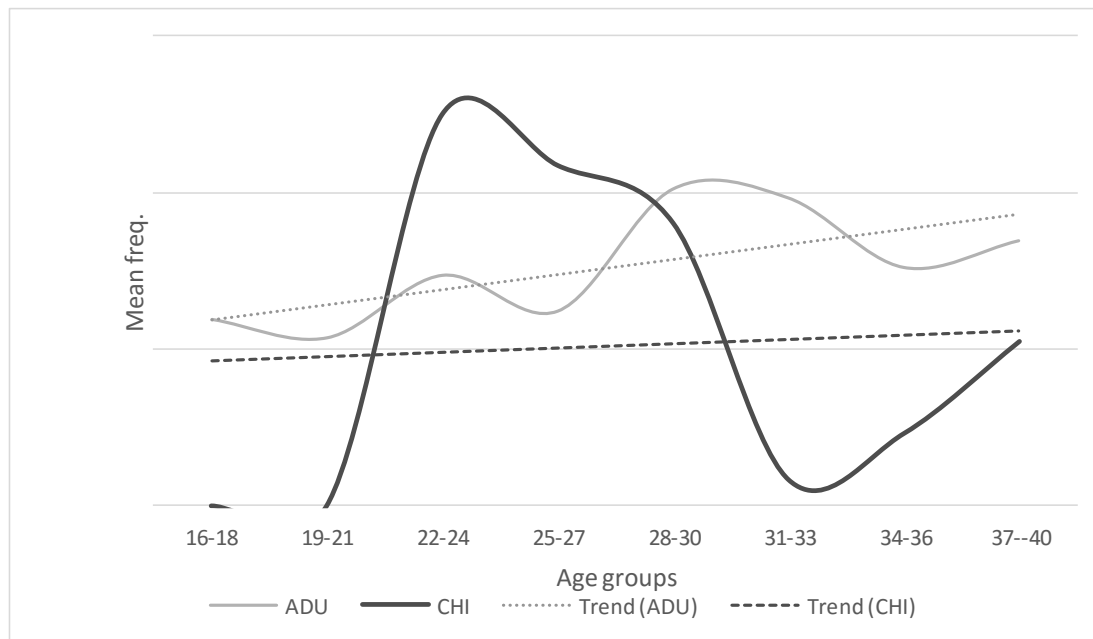

 Figure 9: Distribution of the Perform Transitive Action Construction.

Children produce a few types of this construction with a relatively high token frequency, which indicates that they are rote-learned, entrenched items. This is not surprising, given that the syntax-semantics of this construction is more complex: it involves at least two participants and a causative predicative meaning.

5.4   *The pattern of acquisition*

After studying the distributional trends of the constructions and inspecting the data more deeply, I will now proceed to delineate a tentative general pattern of acquisition for Italian *fare* LVCs.

Clearly, at first children start by imitatively producing those instances that have a relative high frequency in their ambient language: i.e. lexicalized, conventional LVCs. Then, they begin with forming a *fare-* pivot schema in which the free slot is filled with various sound words and is first used to refer to actual sound emission events. Starting from about 24 months of age this schema is then generalized to refer to actions where a volitional agent performs some kind of (intransitive) action. Children indeed seem to use the Perform Sound construction productively to refer to some event or action. (e.g. from *fa pum* 'it emits a pum sound' to *fa pum* 'to shoot at somebody'). Therefore, initially, the schema is applied to refer to simple events with only one participant, which leads to the acquisition of a more general *fare –* ACTION schema (Perform Intransitive Action construction). It is likely that this latter constriction is further extended later to encompass more complex, transitive, actions, leading to the establishment of the Perform Transitive Action construction. [9] Unfortunately the data is insufficient to make a strong claim to this. A dataset covering a longer age period would be needed. Still, given the observation and the trend observed in the data, this is a reasonable hypothesis.

6   Conclusion

This paper focused on the study of the development and use of Italian *fare* LVCs in Child Language and Child-directed Speech as emerging from the analysis of a corpus of naturalistic data.

First, the data was approached assuming the perspective of traditional, formal approaches to LVCs with the objective of showing their weaknesses. The results show that, indeed, there are expressions that bear similarities to typical LVCs, which will not be accounted for according to the traditional definitions. Consequently, I have claimed that they are better accounted for by a constructionist approach to language and specifically as a family of constructions represented as a radial category.

---

[9] It is possible that in adult language only one of the more abstract constructions survives, i.e that the Perform Intransitive and Transitive constructions converge to one abstract fare + EVENT construction. As said above, in order to define the extent, scope and properties of LVCs from a construction grammar perspective precisely, a more extensive and dedicated analysis is needed.

To support this claim, I sketched a potential representation both of the individual constructions and of the whole family. A full specification of the properties of the constructions and their relations in the network is outside the scope of this research and is left for future investigation.

Finally, the analysis of the development of the three most represented *fare* Constructions identified in Child Language led, perhaps, to the most interesting findings of this research: Children's very first productions are indeed rote-learned instances, or fully lexical constructions. The situation is different for each child, of course, but, in general, they use very few types with a relatively high token frequency.[10] Later , they start showing some creativity in their usage. Their first type of LVC is the Perform Sound construction. From 23 months, the increased type variability, in the form of *fare* + onomatopoeic words (e..g. *fare clic* 'make click' for *fare una fotografia* 'take a picture' or *scattare* 'go-off'), can be taken as the first manifestation of a '*fare* +Action/event/situation' pivot schema that enables later combinations with abstract nouns. The use of such a pivot-schema is motivated by the need to name events or situations for which they have no lexical verb yet.

This pattern of development is in line with the predictions and findings of usage-based theories of language acquisition and, in particular, with Tomasello's verb-island hypothesis (1992, 2003): from lexical, item based constructions, a pivot schema is generalized with the function of naming new events, actions or situations. The open slot can be filled first with sound words, then with (abstract) nouns. Interestingly, children seem to learn the general schema starting from a peripheral *fare* construction in adult language: the *Perform Sound* construction.

In spite of some limitations due to the size and density of the corpus, I believe that the results constitute a step toward a full understanding of the development of grammatical constructions in Italian. Far from being simply rote-learned expressions (as they could be if they were unmotivated and fixed or restricted collocations), Light Verb Constructions may instead have an important role in

the acquisition of the more general and abstract transitive construction. Still, many more pieces need to be added to the puzzle.

Future work should be directed towards the analysis of the behavior and acquisition of other light verbs, which would provide us with more information both on the generality/specificity of these constructions and on the way the transitive construction(s) is gradually formed. Secondly, to reach a more comprehensive picture of the acquisition of both LVCs and the transitive construction, a fundamental step appears to be the enlargement of the corpus, possibly with new data based on a denser sampling technique (cfr. Tomasello & Stahl, 2004), and certainly with data from older children. Finally, since any naturalistic observation study bears the intrinsic limitations of studying what there is in the data (cfr. also Theakston et al., 2004), experimental studies along the same line are most welcome.

References

Alba-Salas, J. (2002). Light Verb Constructions in Romance: A Syntactic Analysis. Cornell University. Unpublished PhD Dissertation.

Bates, E. & Goodman, J.C. (1997). On the inseperability of grammar and the lexicon: Evidence from acquisition, aphasia and real time processing. *Language and Cognitive Processes*, 12, 507 – 586.

Camaioni, L. (2001). *Psicologia dello sviluppo del linguaggio*. Bologna: Il Mulino.

Cameron-Faulkner, T., Lieven, E. & Tomasello, M. (2003). A Construction based Analysis of Child Directed Speech. *Cognitive Science* 27: 843-873.

Cantarini, S. (2004). Costrutti Con Verbo Supporto: Una Descrizione contrastiva Italiano - Tedesco. Ferrara: PhD Dissertation. (published as *Costrutti con verbo supporto: italiano e tedesco a confronto.* Bologna: Pàtron).

Croft, W. (2001) *Radical Construction Grammar: Syntactic Theory in Typological Perspective.* Oxford. OUP.

D'Agostino, E. & Elia, A. (1997). Il significato delle frasi: un continuo dalle frasi semplici alle forme poliremathiche. In F. A. Leoni, D. Gamabarara, S. Gensini, F. Lo Piparo & R. Simone (Eds). *Ai limiti del linguaggio*. (pp. 287-310). Bari: Laterza.

De Mauro, T. (2000) *De Mauro. Il dizionario della Lingua Italiana*. Milano: Paravia.

Doğruöz, S. A. & Backus, A. D. (2009). Innovative constructions in Dutch Turkish: An assessment of ongoing contact-induced change. *Bilingualism: Language and Cognition* 12(1): 41–63.

Family, N. (2009). Mapping Semantic Spaces: A constructionist account of the "light verb" eat in Persian. In M. Vanhove (ed.) *From Polysemy to Semantic Change: Towards a Typology of Lexical Semantic Associations*, Amsterdam: John Benjamins.

Fried, M. & Östman, O. (2004). Construction Grammar: A thumbnail sketch. In M. Fried & O. Östman (ed.) *Construction Grammar in a Cross-Language Perspective.* (pp. 11-86). Amsterdam: John Benjamins.

Gaeta, L. (2002). *Quando I verbi compaiono come nomi*. Milano: FrancoAngeli Editore.

Giry-Schneider, J. (1987). *Les Prédicats Nominaux en Français. Les Phrases Simples à Verbe Support*. Geneva: Droz.

Goldberg, A. E. (2006). *Constructions at work*. New York: OUP.

Goldberg, A. E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.

Grimshaw, J. & Mester, A. (1988). Light Verbs and μ-Marking. *Linguistic Inquiry* 19: 205-231.

Iacobini, C. (2000). Base and direction of derivation. In G. Booij, C. Lehmann & J. Mugdan (Eds.) *Morphology. An International Handbook on Inflection and Word Formation*. Vol. 1. (pp. 865-876). Berlin - New York: De Gruyter.

La Fauci, N., & Mirto, I. (2003). *Fare: Elementi di sintassi*. Pisa: Edizioni ETS.

Lakoff, G. (1980). *Metaphors we live by*. Chicago: The University of Chicago Press. (2nd ed. 2008)

Lust, B. (2006). *Child Language: Acquisition and Growth*. NewYork: CUP.

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk*. (3rd Edition) Hillsdale, N. J.: Lawrence Erlbaum Associates.

Miyagawa, S. (1989). Light verbs and the ergative hypothesis. *Linguistic Inquiry* 20: 659-668.

Moon, R. (1998). *Fixed Expressions and Idioms in English: a Corpus-based Approach*, (Oxford Studies in Lexicography and Lexicology), Oxford: Oxford University Press.

Namer, F. (1998). Support Verb Constructions. In F. van Eynde & P.Schmidt (Eds.) *Linguistic specifications for typed feature structure formalisms*. Studies in MT and NLP. Vol. 10. (pp. 315-343). Brussels: CEC.

Ninio, A. (1999). Pathbreaking Verbs in Syntactic Development and the Question of Prototypical Transitivity. *Journal of Child Language* 26, 619-653.

Palancar, E. L. (2003). La polisemia dei verbi dar, pegar e meter in spagnolo. In L. Gaeta & S. Nuraghi (Eds.) *Introduzione alla Linguistica Cognitiva*. Roma: Carocci.

Quochi, V. (2007). A Usage-Based Approach to Light Verb Constructions In Italian: Development and Use. University of Pisa. Unpublished Dissertation.

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2004). Semantic generality, input frequency and the acquisition of syntax. *Journal of Child Language* 31: 61-99.

Tomasello, M. (2003). *Constructing a Language: a Usage-Based Theory of Language Acquisition*. Cambridge, Mass.: Harvard University Press.

Tomasello, M. (1992). *First Verbs: A Case Study of Early Grammatical Development*. Cambridge: CUP.

Tomasello, M. & Stahl, D. (2004). Sampling children's spontaneous speech: How much is enough? *Journal of Child Language* 31, 101-121.