



HAL
open science

Spatial Representation - Experimental and cross-linguistic studies VOLUME 1: CODING MANUAL

Maya Hickmann, Henriëtte Hendriks, Annie-Claude Demagny, Helen Engemann, Anne-Katharina Ochsenbauer Harr, Tatiana Iakovleva, Yingling Ji, Efstathia Soroli

► **To cite this version:**

Maya Hickmann, Henriëtte Hendriks, Annie-Claude Demagny, Helen Engemann, Anne-Katharina Ochsenbauer Harr, et al.. Spatial Representation - Experimental and cross-linguistic studies VOLUME 1: CODING MANUAL. Volume I, 2022. halshs-03836445

HAL Id: halshs-03836445

<https://shs.hal.science/halshs-03836445>

Submitted on 2 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0
International License

Spatial Representation - experimental and cross-linguistic studies

VOLUME 1: CODING MANUAL

Maya HICKMANN
CNRS & Université de Paris 8

Henriëtte HENDRIKS
University of Cambridge

Annie-Claude DEMAGNY
CNRS & Université de Paris 8

Helen ENGEMANN
Universität Mannheim

Anne-Katharina HARR
Ludwig-Maximilians-Universität, München

Tatiana IAKOVLEVA
Université de Paris Diderot

Yinglin JI
University of Shenzhen

Efstathia SOROLI
Université de Lille III

Foreword

This coding manual was developed during a number of cross-linguistic studies regarding the representation and expression of space (the “Langacross” Project) in oral language productions. The four main languages studied initially were English, French, German, and Chinese. In the course of the project, other languages were added (Greek, Russian, Italian, Korean, Tamil), as were other populations. Originally native speakers, child and adult, as well as adult L2 learners were studied. Later bilingual and child L2 data as well as aphasic patient populations were studied using the same materials and theoretical framework.

The work is embedded in a hybrid approach that encompasses formal, functional, and cognitive aspects of language and its acquisition. It takes as its starting point the extensive typological research developed by Talmy (1978, 1985, 2000, 2013) about static and dynamic space. Talmy proposed that the basic motion scene is composed of an entity (the Figure) that is moving in relation to another entity (the Ground), as well as motion per se and the Path of motion. Other types of information (less central to the motion event) such as Manner of motion and Cause of motion are also frequently expressed in the world’s languages. Talmy furthermore proposed that one can distinguish several types of languages, such as those that are *verb-framed* vs. *satellite-framed*, based on how they express the most crucial part of the motion event, i.e., whether they express the Path of motion in the verb root (for verb-framed languages) or in satellites (for satellite-framed languages).

Multiple studies were conducted based on this typology, and it was soon found that the verb-framed / satellite-framed divide is valid particularly when the Path concerns a boundary crossing (cf. Aske 1989, Slobin 2004, and others). Other authors also questioned the possibility of fitting all languages into a bipartite dichotomy. For example, in languages such as Chinese, it is less obvious to define word classes than in other languages because of the absence of morphology, and because verb-like items can be serialised and take on satellite-like features. As a result, such languages might have to be placed in a separate class within Talmy’s typology, which is not clearly satellite-framed, nor clearly verb-framed. Other proposals concern placing languages along a Path- or Manner-salience continuum.

Our studies aim to look in detail into the expression of motion (voluntary and caused motion) across different languages and the impact it may have on language acquisition, including first language acquisition, simultaneous bilinguals, and second language learners at different ages and levels of proficiency. Understanding the impact of the typological differences may bring us closer to the larger research question regarding the relationship between linguistic and cognitive development or, more generally, the relationship between language and thought.

The methodology and coding system presented in this manual allows for such in-depth study of the expression of space (and related temporal information), using formal, functional and semantic categories of two types: 1) general categories that may apply to all languages and 2) language-specific codes that correspond to one particular system. Particular attention is placed on three aspects to be analyzed: what type of information is expressed (information *focus*), how much of it is expressed (semantic *density*), by what particular linguistic means (parts of speech) and/or structures it is expressed (information *locus* and *architecture*).

The manual begins with a general introduction to the data sets that stem from the studies on which the coding system was developed. This system is then described in detail and is organized by coding « fields » (or « levels », as they are called in CHILDES). For each field, the reader will find a general introduction concerning what is coded in it, and why. This part also includes the general coding format for that Field which must be followed strictly if one wants to use tailor-made search programmes that were developed in conjunction with this manual. Following the general introduction, all the different individual codes are provided for this field, with ample examples in English, French and German¹ (information on other languages is provided in the appendices when relevant). Finally, a detailed list of additional remarks, with further explanations on how to code, language- and population-specific codes and warnings about coding pitfalls are presented. It is therefore

¹ Examples in these three languages were not translated, based on the assumption that researchers interested in these languages are proficient in them.

important to always read the full description of a field's coding before embarking on the coding itself.

We wish the users of this manual a lot of coding fun and would appreciate it if they could give us feedback based on their experience with it and/or suggestions for improvements.

References

- Aske, J. (1989) Path predicates in English and Spanish: a closer look. In: Proceedings of the Berkeley Linguistics Society no 15, 1-14.
- Slobin, D.I. (2004) The many ways to search for a frog: linguistic typology and the expression of motion events. In: S. Strömquist & L. Verhoeven (Eds.), *Relating events in narrative: vol.2, typological and contextual perspectives*, 219-257. Mahwah, NJ: Lawrence Erlbaum Associates.
- Talmy, L. (1978) Figure and ground in complex sentences. In: Joseph H. Greenberg (ed.), *Universals of Human Language*, vol. 4: syntax, 625-649. Stanford: Stanford University Press.
- Talmy, L. (1983) How language structures space. In: Herbert L. Pick & Linda P. Acredolo (eds.), *Spatial orientation: Theory, research, and application*, 225-282. New York: Plenum Press.
- Talmy, L. (1985) Lexicalization patterns: semantic structure in lexical forms. In: T. Shopen, R. Keenan, & S. Thompson (eds.), *Language Typology and Syntactic Fieldwork*, vol. 3, 57-149. New York: Cambridge University Press.
- Talmy, L. (2000) *Towards a Cognitive Semantics*. Harvard: MIT Press.

TABLE OF CONTENTS VOLUME 1

Spatial Representation - experimental and cross-linguistic studies	1
1. GENERAL INTRODUCTION	7
1.1 OVERVIEW OF THE TYPES OF CORPORA CONCERNED	8
1.2 TOOLS AVAILABLE	8
1.3 PRELIMINARY STEPS FOR CODING	9
1.3.1 File names	9
1.3.2 Transcriptions	10
1.3.3 Identification of items in the transcription	10
1.3.4 Segmentation	10
1.3.5 Content verification of elicited items	11
1.4 GENERAL CODING PRINCIPLES	11
1.4.1 Coding by picture or 'script' component	11
1.4.2 Use of coding menus	13
1.4.3 General format to follow and summary of coded fields	14
2. EXPERIMENTAL STUDIES	18
2.1 <i>Basic coding lines (%cod)</i>	19
2.1.1 Field 1: Introduction of the coding line	19
2.1.2 Field 2: Component	19
Identifying 'target' components	20
Identifying non-target utterances	25
Important rules	26
2.1.3 Field 3: Response type	28
2.1.4 Field 4: Situation type	29
2.1.5 Field 5: Verb	29
Verb type	30
Verb semantics: information type expressed	32
Distinctions indicated after the infinitives of verbs used (after the §- sign)	33
Prefixation	37
2.1.6 Field 6: Verb Inflection	40
English temporal-aspectual verbal inflection	40
French temporal-aspectual verb inflection	41
German temporal-aspectual verbal inflection	41
Codes relevant for all languages	42
2.1.7 Field 7: 'VOICE'	42
2.1.8 Fields 8 to 12: Satellites	43
Form types (Also see Appendices)	43
Information expressed (see Appendices)	45
2.1.9 Field 13: Periphery	52
2.1.10 Fields 14 and 15: Figures and Grounds	54
Figure and Ground coding	54
Criteria for determining the presence/absence or the relevance of Figure and Ground	56
Additional distinctions	56
2.1.11 Field 16: Comments	59
Common, standardised comments in Field 16	60
Syntax in the bracketed sections of Field 16	61
Annotations that can appear in fields other than Field 16	61
2.2 <i>Synthesized coding lines (%sum)</i>	62
2.2.1 General format of the line	62
2.2.2 Fields 1 and 2: Identification of Item and Component types	62

2.2.3 Field 3: Global structure	63
2.2.4 Fields 4 to 11: Content of Responses	63
<i>2.3 Coding lines for temporality (%tem)</i>	<i>69</i>
2.3.1 General syntax of the line	69
2.3.2 List of Codes in the relevant Fields	69
<i>2.4 Suggestions for the coding of L2 learner and patient data</i>	<i>74</i>
2.4.1 Utterances corresponding to a 'lexical negotiation' (NGL)	74
2.4.2 Code switching (cs)	74
2.4.3 Categorial ambiguities (verb/satellite)	75
2.4.2 Idiosyncratic usages	75
<i>2.5 Information relevant to the coding of Distractor items</i>	<i>78</i>
2.5.1 Introductory remarks	78
2.5.2 Field 2: Item Identification	78
Clauses referring to the Beginning (no change compared to %cod lines)	78
CIB0 Coding lines	79
Clauses referring to the end of a sequence of events in a given item	80
3. SPONTANEOUS PRODUCTIONS	81
3.1 <i>Traditional system</i>	<i>82</i>
3.2 <i>Adaptation of the new system</i>	<i>82</i>

1. GENERAL INTRODUCTION

1.1 OVERVIEW OF THE TYPES OF CORPORA CONCERNED

This coding manual has been compiled based on the coding developed for data collected within the framework of several experimental (controlled production) and longitudinal (spontaneous production with young children at home) studies regarding descriptions of space and expressions of time. The database comprises several languages and several types of speakers (monolingual and bilingual adults and children, L2 learners, and aphasic patients). The manual concentrates on the coding of the three main European languages in the project: English, French and German, but information regarding the coding of other languages is provided in Volume II (Appendices). Codes are either general (language-independent) or they are quite specific and may apply to only one particular language (language-specific, e.g., system of verbal inflections). The coding system was put together for the purpose of analysing the representation of space, and then partly adapted to study the expression of temporality. It is based on and therefore mainly suitable for two types of corpora:

- (1) productions collected in controlled experimental situations with different groups of people (monolingual or bilingual children and adults, L2 learners, aphasic patients, etc.);
- (2) spontaneous productions collected within the framework of longitudinal studies with young children at home.

Several tasks have been employed for the experimental studies, each of which uses different items to look at a particular aspect of spatial representation. The examples in the manual were extracted from the following tasks (but the coding and hence future manual may possibly be adapted to other tasks):

- Voluntary motion 1 (*Voluntary motion cartoons*: ‘up, down, across’).
- Voluntary motion 2 (*Coralie cartoons*: ‘up, down, across, along, into, out of’).
- Voluntary motion 3 (*Eva videos*: ‘up, down, across, along, into, out of’).
- Caused motion (*Hoppy cartoons*: ‘push, pull, roll, slide, up, down, into, across’).
- Narrations (two picture sequences ‘*Cat*’, ‘*Horse*’ (Hickmann, 2003) and one picture book ‘*Frog where are you*’ by Mercer Mayer 1969).

1.2 TOOLS AVAILABLE

When coding our data, we made extensive use of the tools available in CHILDES (2013) and we therefore formatted the files in a compatible CHAT format. The tools in CHILDES enable the effective execution of analyses for which this manual has been developed. However, the coding principles presented here can also be adapted to other systems.

CHILDES in general:

A CHILDES manual is available online. For the most relevant information see the CLAN Manual (<https://talkbank.org/manuals/CLAN.pdf>) and in particular the sub-section on the Coder Mode in the Editor chapter (codes file, n-level taxonomic hierarchy) and Chapter 7 which introduces the analysis tools.

For coding:

CHILDES enables the production of specific files offering coder menus to enable the codes to be entered semi-automatically. Files created in this project and offering ‘menus’ of codes adapted for our particular coding are also available. Some of our codes apply to all languages, but others are specific to each language. The coding file names therefore all begin with ‘codexp’, but are then followed by a three-letter abbreviation for the language in question (e.g., codexpgerspace.cut for German).

For file checking after coding, and for analysis:

To ensure the proper functioning of all CHILDES and other programmes for analysis, it is essential to run the coded files through the CHILDES CHECK programme, analyse the coding lines field by field with *FREQ*, and use this programme to correct errors (cf. CHILDES manual link above). Some of the CLAN commands can

also be used for the analyses (particularly **FREQ** and **COMBO**).

Other programmes for analyses: Some programmes have been specifically written by our research team for the analysis of data coded with the coding system described in this manual. Two programmes are used to analyse the coding of spatial expressions on the %cod lines and %sum lines in the experimental data (**resesp_new.exe** and **synth.exe** respectively).

1.3 PRELIMINARY STEPS FOR CODING

1.3.1 File names

When creating a large database, we recommend deciding early on the naming of individual files. For our data, all files were named as follows:

- Initial(s) indicating the language: C Chinese, E English, F French, G German, Gr Greek, I Italian, etc.
- Age group in numbers for children (i.e., 06 for 6-year-olds), AD for adults.
- Subject number (i.e., 0601 for 1st participant in the 6-year-old group).
- Task identification (vol, pop, cat, hor, fro, etc.), followed by the file type.

Once the coding is completed, the files fall into two types, **ori** and **tot**:

ori: original transcribed files (no coding). **...popori**
tot: completely coded file (all relevant coding lines finished). **...poptot**

In order to keep track of coding progress, several temporary files can also be created, the names of which indicate which lines have been coded and whether or not they have been checked. Such detailed naming may be particularly useful when multiple coders are likely to work on the same files whilst executing different tasks. The naming used in our project is as below:

- **cod:** %cod lines coded.
Basic spatial coding.
- **sum:** %sum lines coded.
Synthetic spatial coding taking account of all information in a given response.
- **tem:** %tem lines coded.
Basic coding by utterance more specifically related to temporality.
- **sut:** %sut lines coded
Synthetic coding more specifically related to temporality.

Examples:

- E0401popori:
English speaking four-year-old child, 1st subject in this age group, ‘Hoppy’ task (French ‘Popi’), original transcribed file.
- F0401popcod:
French speaking four-year-old child, 1st subject in this age group, ‘Hoppy’ task, coded for %cod lines.
- GAD01poptot:
German speaking adult, 1st subject in age group, ‘Hoppy’ task, file coded with %cod and %sum for spatial lines.

- C0803voltot:

Chinese speaking eight-year-old child, 3rd subject in this age group, ‘Voluntary motion’ task, file fully coded.

1.3.2 Transcriptions

The productions to be coded are transcribed according to the specific format set out by CHILDES (see in particular Chapter 5 of the CHILDES CHAT manual <https://talkbank.org/manuals/CHAT.pdf>). In order for the CLAN programmes and those provided by the current research project to work, it is essential, among other things, that tabulation [tab] is used, and that no spaces are inserted in the coding lines. Before coding, ensure that the file is correctly and completely transcribed. It should run through CHECK without any problems detected. For elicitation materials that comprise multiple items, the items must be in the correct order and the utterances segmented according to the instructions provided below.

1.3.3 Identification of items in the transcription

In the case of productions elicited by means of a series of items in an experimental situation, responses are identified corresponding to each item in the transcription. A line starting with the @G sign, followed by a colon and a tab and then the item name, is inserted before the corresponding transcription of the item. For this purpose, CLAN files specifically produced for our materials and containing the lists of items in order of presentation and corresponding to each entry in each task (‘skeleton’ files) are available, thereby enabling the transcriber to avoid having to retype the item names for each subject (See Appendix V for an example). The complete set of item names is given in Appendix X in Volume II).

@G: [tab]prt_pourue (experimental item combining pushing, rolling and going across (traverser))

*SUI: [tab]Hoppy pushes a pram across the road.

@G: [tab]tge_chagro

*SUI: [tab]alors maintenant il entre dans la grotte en tirant une chaise derrière lui.

@G: [tab]XVEL (experimental item “vélo”)

*SUI: The girl rides her bike across the railway track

@G: [tab]CBEB (control item “bébé”)

*SUI: The baby is crawling

In the case of (Cat, Horse, and Frog) narratives where item names are irrelevant, names are replaced by the identification of the pictures being referred to in each segment of the production. This identification always begins with the ‘pict_’ sign followed by the number of the picture (i.e., pict_01). Special codes are also used when it is ambiguous what picture is being referred to. For more detailed information, see section 3.1 below.

Example:

@G: [tab]pict_01

*SUI: [tab]There's a tree with some birds in a nest.

@G: [tab]pict_05

*SUI: [tab]The dog pulls the cat down by the tail.

1.3.4 Segmentation

Each segment must contain a verb allowing for certain exceptions. The verb can be finite or non-finite in the case of certain subordinate clauses. For example, gerunds such as ‘*en courant*’ and infinitives such as ‘to leave’ are treated as belonging to distinct clauses.

Exceptions (see detailed description of Field 5 for additional information):

- 1) The utterance contains a complex verbal expression involving more than one verb, but is treated as a single clause, and corresponds to one coding line only. For example, in English, German or French, modals (*want, can, vouloir, pouvoir, wollen, können*) and temporal-aspectual constructions (*begin to, finish, être en train de, se mettre à, anfangen*) are not segmented in the coding.
- 2) Some utterances do not contain explicit verbs, because the verb has been omitted through ellipsis or anaphora, for example, presentative clauses containing labelling without a verb (*‘Un nageur. Il plonge...’*). These cases are considered as representing two clauses, even though the first one does not contain a verb.

Both for narrative and individual item data, segmentation is marked by adding a [c] after each clause, each clause corresponding to one coding line². If the utterance includes just one clause, it will only contain one symbol [c] at the end. As an example of this coding of the segmentation, the following utterance includes three clauses and must therefore be followed by three lines of coding:

*SUJ: and there's a little boy [c] sliding across the ice [c] and running back onto the snow [c].

In the case of individually elicited items in an experimental set-up (voluntary motion, hoppy), the response to a given item may consist of several utterances (not all of which are necessarily relevant for the research question). All output (relevant or not) is segmented into clauses. Each clause will correspond to one coding line. These coding lines are located immediately after the utterance to be coded, i.e., after the *SUJ line, in the order in which the clauses are uttered.

1.3.5 Content verification of elicited items

If the content of the utterance to be coded does not correspond to the item (different locations, referents or objects mentioned), this can be due to an error in the item name, an omission, or the item being inserted incorrectly in the file. In this case, and following verification, the transcription should be corrected or completed before coding. The ‘skeleton files’ mentioned in 2.3 above are useful for this verification.

1.4 GENERAL CODING PRINCIPLES

Once transcription and segmentation are finalised, and content has been checked, coding can begin. For purposes of automatic analyses, coding lines should always follow immediately after the *SUJ lines, as shown in this manual. When intervening *EXP lines precede the coding lines, the coded information will not be considered by the automatized analyses.

1.4.1 Coding by picture or ‘script’ component

For picture supported narratives (Cat, Horse, Frog):

All the utterances are analysed, and the coding is carried out by clause. The coding contains an indication of the pictures being referred to.

*SUJ: *Y a un chat [c] qui arrive [c] et il s'assied sous l'arbre [c].*
%cod: \$:pict2:...Coding for *Y a un chat*
 \$:pict2:...Coding for *qui arrive*
 \$:pict3:...Coding for *et il s'assied sous l'arbre.*

For other experimental data:

The coding is centred around those utterances corresponding to the intended targets in the experimental plan

² For further information on the segmentation of narratives, see Hickmann, Hendriks, Roland and Liang (1994). The Development of Reference to Person, Time and Space in Discourse: A Coding Manual.

is facilitated with the activation of TEXT Mode (Esc m). The other way of alternating between the two modes is by accessing the “mode” tab in CHAT.

1.4.3 General format to follow and summary of coded fields

a) Fields:

Each coding line comprises fields, some of which may comprise sub-fields. Fields and sub-fields are separated by determiners. Fields are coded in a set order, and all follow a set format. Irrespective of the type of coding line, fields are always separated by a colon (:). Sub-fields are separated by various signs (#, &, |, \$).

There are four different types of coding lines, each beginning with a different symbol:

- **%cod**: main coding lines, corresponding to items that are central to the study;
- **%tem**: coding lines for temporality;
- **%sum**: synthetic coding lines;
- **%sut**: synthetic temporality coding lines.

Each line comprises the following number of fields:

- **%cod**: 16 fields;
- **%tem**: 16 fields;
- **%sum** and **%sut**: 10 fields.

Example of a %cod complete coding line for the target information:

*SUI: **a monkey climbed up the tree** [c] and got a banana [c].

%cod: \$:CIBM:LR01:SD:VMM#+Mmv|climb:PAS:DAC:EP1/1&+Tve|up+NP:EZ2/1&|:
EZ3/1&|EZ4/1&|EZ5/1&|PM0:Fig|monkey:Grd|tree:{ }

b) Summary of fields:

The lists below provide a summary of the fields for each type of coding line.

%cod lines:

- (1) Introduction of the coding line (\$).
- (2) Component (COMP): Beginning, End, Target or possible Target; Picture.
- (3) Response type (RESP): spontaneous or induced, numbering.
- (4) Situation type (SIT): static, dynamic or indeterminate.
- (5) Verb type and information expressed by the verb (VERB), followed by the infinitive, the quotation form and/or the most current form of the verb.
- (6) Verbal inflection (INFL): varies for each language, to be adjusted for each group of learners.
- (7) Diathesis (VOICE): active, passive, medium.
- (8-12) Satellites (SAT): form type (preposition, particle, adverb, etc.), followed by information expressed by the satellite, then the specific form used.
- (13) Periphery (PERIPH): identification of independent main (matrix) clauses and subordinate clauses.
- (14) Figure (FIG): indication of the referent functioning as the Figure.
- (15) Ground (GR): indication of the referent functioning as the Ground.
- (16) Comments may follow in brackets { }.

%tem lines:

- (1) Introduction of the line (\$).
- (2) Component (COMP): Beginning, Target, End, or picture.
- (3) Response type (RESP): spontaneous or induced, numbering.
- (4) Predicate typology (PREDTYP): Klein, Vendler, Durativity, Boundedness.
Verb type and information expressed by the verb (VERB) followed by the verb infinitive.

- (5) Verbal morphology (INFL): Inflection used (varies for each language, to be adjusted for each group of participants).
- (6) Verbal inflection breakdown: Tense.
- (7) Verbal inflection breakdown: Aspect.
- (9-13) Satellites (SAT): form type (preposition, particle, adverb, etc.), followed by information expressed by the satellite, then the specific form used.
- (14) Periphery (PERIPH): independent main (matrix) clauses and subordinate clauses.
- (15) Figure (FIG): indication of the referent functioning as the Figure.
- (16) Ground (GR): indication of the referent functioning as the Ground.
- (17) Comments may follow in brackets { }.

%sum lines:

- (1) Introduction of the line (\$).
- (2) Identification of the item.
- (3) Global structure of the response (relative complexity and nature that is more or less dense).
- (4-11) Information relating to each component, each comprising three sub-fields.

%sut lines:

These lines are identical to the %sum lines, additionally comprising information relating to temporality, in particular the marking of temporal boundaries.

c) Syntax of fields:

%cod lines:

%cod:[tab]\$:COMP:RESP:SIT:VERB#+sem+sem|infinitive\$-predicate:INFL:VOICE:
 SAT1/N&+sem+sem|form:SAT2/N&+sem+sem|form:SAT3/N&+sem+sem|form:
 SAT4/N&+sem+sem|form:SAT5/N&+sem+sem|form:PERIPH:FIG|form:
 GR|form:{ comments }

%sum lines:

%sum:[tab]\$:item_name:GLOBAL_{tot}:COMP_a:sem&sem#form\$-form:COMP_b:sem&sem#
 form\$-form:COMP_c:sem&sem#form\$-form:COMP_d:sem&sem#form\$-form

%tem lines:

%tem:[tab]\$:COMP:RESP:Klein#Vendler#durativity#boundedness\$-type-of-verb-or-type-of-satellite-or-
 z:VERB#+sem+sem|infinitive\$-predicate:INFL:Tense:Aspect:
 SAT1/N&+sem+sem|form:SAT2/N&+sem+sem|form:SAT3/N&+sem+sem|form:
 SAT4/N&+sem+sem|form:SAT5/N&+sem+sem|form:PERIPH+sem|form:FIG|form:
 GR|form:{ }

%sut lines:

%sut:[tab]\$:item_name:GLOBAL_{tot}:COMP+boundary_a:sem+boundary&sem+boundary#form\$
 -form:COMP+boundary_b:sem+boundary&sem+boundary#form\$-
 form:COMP+boundary_c:sem+boundary&sem+boundary#form\$-form:
 COMP+boundary_d:sem+boundary&sem+boundary#form\$-form

Example (1) below contains three clauses, including two target clauses expressing the up and down movement and a clause that contains information irrelevant for motion analyses. No BEG or RSF information is provided in this response, leading to the coding as indicated.

*S_{UJ}: *the squirrel climbed up the tree [c], went in the hole [c] and climbed back down [c].*
 %cod: \$:BEG/NO:0:0:0#0|0:0:0:0&0|0:0&0|0:0&0|0:0&0|0:0&0|0:0:Fi0|0:Gr0|0:{ 0 }

paragraph and hyphen (§-)

delimiter located after the verb infinitive and/or after the form of the satellite, followed by other distinctions.

...jump§-want (to code *he wants to jump*)

...jump§-neg (to code *he doesn't jump*)

hash (#)

delimiter indicating the beginning of semantic information expressed in the verb.

ampersand (&)

delimiter indicating the beginning of semantic information expressed in the satellites.

hyphen (-)

connects elements on the coding line that go together in complex expressions:

Example: *From-left-to-right, faire-courses*

plus (+)

used in front of every semantic component coded, irrespective of whether the information occurs on its own (for example, +Mmv for *walk*), or in combination with other information (+Tve+Mmv for *grimper*).

used to indicate that more than one referent is mentioned as Figure or Ground in the item (e.g., *he is walking on the pavement towards the road* pavement+road).

Used inside { } brackets: when several comments listed.

equals (=)

Can be used in the final “comment” brackets to indicate equivalence, especially when glosses are used with use of linguistic idiosyncracies.

2. EXPERIMENTAL STUDIES

2.1 Basic coding lines (%cod)

In the following we will present information on the coding system by field and as the fields occur on the %cod line. Please note that below information in the following three sections (2A-2C) is only suitable for the coding of the experimental items.³ For the full set of training, experimental and distractor items for our main datasets, i.e., *Hoppy* and *voluntary motion cartoons*, please refer to Appendix x in Volume 2.

2.1.1 Field 1: Introduction of the coding line

Each line starts with the symbol \$.

2.1.2 Field 2: Component

This field codes utterances produced as components linked to each task / item. Each task therefore has specific codes.

1. For the narratives (Cat, Horse, Frog):

All utterances are analysed and the coding in this field relates to the picture being referred to. Depending on which specific picture is indicated in the transcription, the following codes are used in Field 2:

- Clear reference: pict1, pict2...
- Ambiguous reference: px2 (not entirely sure if reference to picture 2)
- Inferred from a picture: pict_1
- 'Meta-story' utterance: pict_0
- Utterance unrelated to story: pict_0

*SUJ: *This is a story about a cat [c]. There's a tree with some birds in a nest [c]. There comes a cat [c], and it watches the birds [c]. It climbs up [c]...*

%cod: \$:pict0:...Coding for *This is a story about a cat*.
\$:pict1:...Coding for *There's a tree with some birds in a nest*.
\$:pict2:...Coding for *There comes a cat*
\$:px2:...Coding for *and it watches the birds (could refer to pictures 2 or 3)*
\$:pict4:...Coding for *It climbs up*

2. For other experimental data - voluntary and caused motion events:

The coding is centred around those utterances corresponding to the intended targets ('Cibles') in the experimental design, but also includes other information relating to the rest of the output. The script corresponding to each item is analysed in several 'basic' components, including:

- A. The target (cible) or possible targets (mainly those utterances being sought after for the analysis);
- B. Other utterances accompanying the target.

➤ Target or potential targets:

CI... and MI... All utterances that are treated as targets or potential targets in the coding are identified by a

³ Some of our experiments included distractor items (Hoppy experiment). Such items differed from experimental items in that the force provoking the motion of an entity was inanimate (e.g. a ball rolling into a book that then slides) rather than animate (Hoppy). The coding and analysis of these items are in progress. These items require a specific coding procedure which will be described in a separate section below (section 2D).

four-character code (see below for details):

This code always begins with the two letters **CI** (cible = target) or **MI** (possible target).
A third letter in this field provides additional coding information:

CIB: if the identification of the target is straightforward, the third letter is always B;

CIR, CIP, CIA, CIS, CIX: if the target coding requires the coder to select from amongst a number of possible targets, the third letter (R, P, A, S, X) indicates the criterion that has been used to select the target;

MIR, MIP, MIA, MIS, MIX: when a target has been selected, these codes identify those utterances that have not been selected. They always begin with MI, and the same third letter indicating the criterion used to select the target.

- The fourth character is used to differentiate the type of target in various items: The ‘up/down’ items in the Voluntary motion data comprise two codes (M=up (monter), D=down (descendre)), and other items have one code (the number 0):

CIBM, CIRM, CIPM, CIAM, CISM, CIXM for ‘up’ situations (‘up/down’ items).

CIBD, CIRD, CIPD, CIAD, CISB, CIXD for ‘down’ situations (‘up/down’ items).

CIB0, CIR0, CIP0, CIA0, CIS0, CIX0 (other items: ‘across’, ‘Hoppy’).

- And their equivalents in those cases where one target utterance is selected from several possible options, for example:

MIRM, MIPM, MIAM, MISM, MIXM for ‘up’.

➤ Other utterances:

BEG Beginning of the script (setting the scene, initial locations, etc.).

RSF Resulting or final state, or departure from scene.

IRR Irrelevant utterances, those referring to elements that are not relevant for the coding of motion as not targeted (‘irrelevant’).

Identifying ‘target’ components

Some statistical analyses require one target response only per subject and per item to be selected, which may comprise several coding lines (in the case of a matrix clause with subordinate clauses). This constraint can be ignored in the qualitative analyses, and in our system a supplementary coding line (the synthetic line) will take account of the way the information has been distributed between the utterances in the discourse (see Section 2B).

Before coding the target field, it is important to understand what information is being sought as ‘target’ for the different items. The main principle for statistical analysis is to find the utterances that express the largest quantity of relevant information, which will vary depending on the items:

a) Voluntary motion cartoons

The voluntary motion cartoons include two types of items, i.e., experimental and control items. The experimental items show agents moving in a specific Manner and with regards to clearly defined Paths. Control items downplay Path by showing agents moving in a specific Manner against an empty grey background. Target components differ depending on the items:

- Experimental items:

Utterances simultaneously expressing both Path and Manner shown in the item. The main Paths for these items are the upward and downward Paths, and the ‘across’ Path.

- Control items:

For control items, only Manner is expected in the responses (and not Path). Manner is thus the targeted information.

b) Caused Motion (Hoppy)

Utterances that simultaneously express Cause and Agent Action, as well as Path and Manner (relative to the Object) as shown in the item.

- Cause and Agent Action: *push, pull*
- Path of the Object: *up, down, across, and into*
- Manner of the Object: *slide, roll*

Note: If a subject uses *pull* instead of *push* or *down* instead of *up*, and no other utterance in the output describes the target event, do not code IRR, but indicate VMM#+Cse+Mac|pull\$-intrans-**idiouv-push**.

To support the coders in finding the target events, further information has been provided in Volume 2, Chapter 2 on Targets.

In many cases, the utterances produced do not contain all the information expected or contain unexpected information. Furthermore, the coding of target utterances depends to some extent on the complexity of the response. In some ‘simple’ cases, the subject only produces one utterance that can be identified as target (CIB0, CIBM, CIBD). In other more ‘complex’ cases, the subject produces several utterances that could be understood to refer to the target events. In such cases, one needs to be chosen and indicated by CI..., while the others that have not been chosen will be indicated by MI...

Several criteria are used in a hierarchical fashion to make these decisions: the ‘richness’ criterion is applied first, followed by several other criteria, which may also be applied hierarchically. Finally, some cases may present coding difficulties that cannot be resolved. In our data, we found that around 95% of the cases allowed for a clear identification of the Target, either because participants only provided one utterance referring to the target event in their response, or because there was a clear richness distinction. Targets identified by other criteria as below are therefore very infrequent.

The full set of codes for this field is summarised in Table 1, below. A more detailed description of the various coded follows after the Table.

Table 1. Summary of codes in Field 2 with regard to targets and potential targets

CIB...: Simple cases

Only one utterance can be considered the target (coded CIB...).

CIR.../MIR...: First criterion for complex cases: ‘Richness’

Two or more target utterances are possible, but one of them (coded CIR...) is richer in semantic content than the others (coded MIR...). For how to determine richness, please see below.

CIP.../MIP...: Second criterion for complex cases: ‘Primacy’

Two or more utterances are equally rich, but one of them contains information that is more important for the item in question, (coded CIP...) than the others (coded MIP...).

CIA.../MIA...: Third criterion for complex cases: ‘Agentivity’

Two or more utterances are equally rich and express equally important parts of information for the item. However, only one of them concerns the agent (coded CIA...), whereas the others concern the object (coded MIA...).

CIS.../MIS...: Fourth criterion for complex cases: ‘Subjective’ relevance

Two or more utterances are equally rich and are also equal in terms of the Primacy and Agentivity criterion. However, one is clearly more pertinent for the experiment (coded CIS...) than the other (coded MIS...). This subjective criterion should be used with caution.

(CIQ...)/MIQ...: Fifth criterion for complex cases: Specific ‘Questions’

In cases of utterances in response to a specific question:

- If there is only one possible target, and that target utterance is given in response to a specific question, code CIQ.
- If there are two possible targets, of which one is in response to a specific question, code the response as MIQ, and the other possible target as CIB.

MID... Sixth criterion for complex cases: ‘Doubles’

The more or less exact repetition of an utterance. If this results in two targets (the target and its repetition), code: CIB and MID. In more complex cases, code: CIR (target), MIR (potential target, but less rich) and MID (the double of either the CIR or MIR coded utterance).

CIX.../MIX...: Coding difficulties

Cases in which it is impossible to identify a target based on all of the above criteria.

1. 'Richness' criterion

In some cases, several target utterances are possible, but one amongst them is richer in terms of semantic content. To determine the relative richness of various utterances, the number of different types of semantic information expressed (not sub-types) should be taken into account. For example, if an utterance contains several pieces of information regarding the Path, such as the source and goal, and another clause expresses information regarding Path and Manner of motion or Cause and Manner of Cause, the second utterance is selected, as this contains a larger variation of types of semantic information. In the examples below, the target utterances are indicated in bold (see Fields 5 and 8-12 below for the coding of Path (T), Manner (M), Cause (C) and Location (L)):

- **Hoppy runs down** [CIR0=M+T], *he's still going down* [MIR0=T].
- **The worm crawls up** [CIRM=M+T], *he is still running* [MIRM=M].
- *And he goes all the way down* [MIRD=T], **he runs down** [CIRD=M+T].
- *He is going down the hill* [MIR0=T], **he is rolling the ball** [CIR0=M+C].
- **Le chat monte en courant** [CIRM=T+M], *il court encore* [MIRM=M].
- *Il traverse d'un côté jusqu'à l'autre* [MIR0=T+T+T] **et pousse le panier** [CIR0=C+M].
- *Il monte la dune* [MIR0=T] **et pousse la valise** [CIR0=C+M].
- *Popi geht einen Berg hinunter* [MIRD=T+T] **und zieht dabei eine Schubkarre hinter sich her** [CIRD=C+M, L+T].

2. Other criteria (presented hierarchically)

Several target utterances are possible, and at least two of them are equally rich. In these cases, the choice of target utterance is made according to the following hierarchy of criteria: 'Primacy', 'Agentivity', and 'Subjectivity'.

Primacy criterion:

The Path is considered as the prime semantic information for items in which both Path and Manner are equally salient (experimental items). Manner is considered as the prime semantic information for items in which only Manner is salient (control items in voluntary motion materials).

- Experimental item: *A baby. Oh no, it's going across the road* (CIP0), *it's crawling* (MIP0).
Control item: *L'ours marche* (CIP0) *et il avance* (MIP0).

Agentivity criterion:

Utterances that refer to the agent have priority. This may happen when utterances express the Manner of the agent (Mma) versus the Manner of the object (Mmp); utterances expressing the agent's Path versus the object's Path; or when active (focus=agent) versus passive utterances are used in response to the same item.

- **Hoppy is walking** and *the ball is rolling*.
- **Popi descend** et *la balle va par là*.
- **Hoppy pulls the car up**, *the car is pulled up*.

Subjective Relevance:

One utterance is clearly more relevant than the other(s) in the eye of the coder. This criterion is to be used with care, and only if the coder is sure not to be ignoring the criteria explained above (richness, primacy, and/or agentivity).

- **It's going down** (CIS0) *and it goes all the way to the bottom* (MIS0).
- *Il descend du trottoir* (MIS0) **et il traverse la route** (CIS0).

NB. If the participant in the latter case had just said *Il descend du trottoir*, the utterance would be coded as target.

3. Special cases

CIQ/MIQ: Preference is given to spontaneous responses or responses to general questions, but not to responses to specific questions (which are coded with MIQ). The MIQ code can go with any CI... or other MI... codes.

- a) If there is only one possible target, and that target utterance is given in response to a specific question, code CIQ.

*EXP: now a mouse, there's a mouse and a table and some cheese.
*SUI: he eats it. (IRR)
*EXP: all right, what happened then? (QG)
*SUI: he was getting it. (IRR)
*EXP: all right, he was getting the cheese, how did he get the cheese? (QS)
*SUI: climbed up the table. (CIQM)

- b) If there are two possible targets, one of which is in response to a specific question, code the response as: CIB and MIQ.

*SUI: Il monte. (CIB0)
*EXP: Comment est-ce qu'il a monté la balle? (QS)
*SUI: Il l'a roulée. (MIQ0)

MID...:

- a) The more or less exact repetition of a response already given. In this case, the MID code can go with any CI... or other MI... codes. If this results in two potential targets (a target and its repetition), code: CIB and MID.

- *He's running up fast (CIBM) ... he's running up at top speed... (MIDM)*

If it results in three targets, code CIR (target), MIR (potential target but less rich) and MID (the double of either the CI or MI coded utterance).

- *Il monte en courant (CIRM) ... il monte (MIRM) Il est monté avec ses pattes (MIDM).*

- b) If one of the utterances is negated (neg), choose the positive one.

- *Der Mann läuft nicht (MID0) **er rennt** (CIB0).*

4. 'Unresolved' cases

Cases in which a target based on all the above criteria cannot be identified. The choice of the target is arbitrary (in principle, these cases are very rare).

- *He went from the left side to the right (MIX0) **and he went across the road** (CIX0).*
- *Il nage jusqu'au bout (MIX0), **il traverse la rivière en crawl** (CIX0)...*

5. Additional codes

The following code may be added in brackets at the end of some lines beginning with MI... (MIR..., MIP..., MIS..., MIX...):

{**repetition**} indicates that a MI... line is a repetition (more or less the same, but very similar) to another CI... line. Example:

- *And he slides down the table (MIDD =Mmv+Tve) **He slides down the table leg** (CIBD =Mmv+Tve)*

- *Il monte la bouée* (CIR0 =Cse+Tve)... *Il monte la bouée bleue* (MID0=Cse+Tve).

Identifying non-target utterances

1. Components corresponding to ‘beginning’ (BEG)

There can be several coding lines starting with BEG. These utterances provide information relating to the setting (situations, locations). They often appear at the beginning of the story, but not necessarily. They are often, but not always, static, and are made up of examples with verbs such as *existential verbs*, *partir de*, *arriver*, and *sich befinden*. Examples:

- *A mouse is tiptoeing up to a table* (=BEG), *now he's climbing up the table leg* (=CIBM) and *about to nick the cheese* (=IRR).
- *Hoppy had the shopping trolley* (=BEG) and *brought it into the cave* (=CIB0).
- *Il arrive* (=BEG), *il descend* (=CIBD), *il y avait une colline de neige* (=BEG) et *il sourit* (= IRR).
- *Der junge ging den Berg hinunter* (=BEG), *durch das Wasser schwamm er* (=CIB0).
- *Popi befindet sich vor einem Bauernhaus mit einem schrägen Dach* (=BEG) und *schiebt ein großes lilanes Geschenk vor sich das Dach von ganz unten bis ganz oben an den First* (=CIBM)

2. Component corresponding to the ‘end’ (RSF or ‘final / resulting outcome’)

There can be several coding lines starting with RSF. These utterances generally appear at the end of the story. Utterances containing verbs such as *arrive*, *atteindre* or *be back at the bottom*, *se retrouver sur l'autre trottoir*, *davonlaufen*, *verschwinden* are often, but not always, of this type. Examples:

- *Das eichhörnchen hüpfte zwischen den bäumen* (=BEG) und *klettert an einem stamm hinauf* (=CIBM) und *verschwindet in einem astloch* (=IRR) *kommt wieder heraus* (=IRR) *klettert wieder herunter* (=CIBD) und *verschwindet wieder* (=RSF).
- *Hoppy leaves from the left* (=BEG), *crosses the road* (=CIB0) and *reaches the other pavement* (=RSF).

But: Hoppy crosses the road (=CIB0) to arrive on the other pavement (=CIB0)

When accompanied by a clear target, some verbs are used particularly frequently in the initial part of the scenario (BEG, for example, *come*, *arriver*, *ankommen*) or in the end part (RSF, for example, *reach/atteindre*). However, they should only be treated as targets in the rare cases where there is no other utterance that can be considered as such.

3. Irrelevant utterances (IRR)

The code ‘IRR’ is used for utterances or clauses that are not part of any of the above components. This code enables these utterances to be spotted, which is particularly important when they relate to clauses that are considered relevant. Coders may decide to code these lines or not. In the latter case, all fields will comprise zeros. Examples:

- *the caterpillar is climbing up the plant* [coded] *about to munch on a leaf* [IRR] and *then he's climbing back down again* [coded].
- *a cat ran up a lamp post* [coded], *then licked away an egg after dropping it from a nest* [IRR] and *ran off of the screen* [coded].
- *L'écureuil monte* [coded], *il rentre* [IRR], *et sort de son trou* [IRR].
- *Il arrive en bas* [coded]. *Il est content* [IRR], *et il sourit* [IRR].
- *Der affe ist in die palme gegangen* [coded], *hat eine banane genommen* [IRR] und *ist dann wieder herüber gegangen* [coded].

\$:IRR:0:0:0#0|0:0:0:0&0|0:0&0|0:0&0|0:0&0|0:0&0|0:0:Fi0|0:Gr0|0:{0}

Important rules

1. A single target utterance

For each target displacement, only one utterance coded with the 'target' code (i.e., with the letters CI...) is required, whether this is a clear target (CIB...) or a less clear target (CIR..., CIP..., CIS..., CIX...). The target utterance may result in several CI coding lines, i.e., when it includes subordinate clauses.

2. Coding pairs in 'complex' cases

The coding pairs used for coding difficulties mostly go together, with only a few exceptions:

- if an utterance is coded CIR..., it must have at least one complementary utterance coded MIR...
- if an utterance is coded CIP..., it must have at least one complementary utterance coded MIP...
- if an utterance is coded CIS..., it must have at least one complementary utterance coded MIS...
- if an utterance is coded CIX..., it must have at least one complementary utterance coded MIX...

Exceptions:

The codes MIQ and MID can be used with all CI... or MI... codes.

3. 'Mixed' coding lines

If an utterance is coded as a clear target (CIB0, CIBM, CIBD), no mixed type (i.e., no MI... code) utterance should be coded. However, where a less clear target (codes beginning with CIR, CIP, CIS, CIX) is involved, there can be one or more additional mixed utterances. For example, for one code beginning with CIR, there can be several utterances beginning with MIR.

4. Problems linked to subordination (see also field 13)

(a) The target can be a subordinate clause connected to a main clause expressing another component. In this case, the periphery is coded with the special code reserved for subordinate clauses that can be treated as main clauses (PMX, see field 13 for details). Examples:

- *this time it's a grey kitten (BEG) which runs quite quickly across the middle of the screen from left to right (PMX target).*
- *il est malin (IRR) parce qu'il est monté (PMX target).*
- *der bär kommt von rechts (BEG), ziemlich schnell laufend (PMX target).*

(b) Subordination is a criterion that may influence what information is to be included in the target clause. In this way, some information that is relevant to the beginning (BEG) or the end (RSF) can be included in the target if it corresponds to subordinate phrases in a larger structure that includes the target. Examples:

- *the woman riding her bike [CIB0] gets to the other side of the screen [CIB0].*
but: *the woman rides her bike [CIB0] and gets to the other side of the screen [RSF].*
- *Ein kleiner junge nimmt anlauf [CIB0], um über einen gefrorenen fluss zu rutschen [CIB0].*
but: *Ein kleiner junge nimmt anlauf [BEG] und rutscht über einen gefrorenen fluss [CIB0].*

5. 'Forced' choice in the absence of a target

In certain cases, independent clauses (without subordination) which can be considered beginning or end utterances, or IRR ('irrelevant'), are nevertheless taken into account and coded as targets if they have some relevance to the description of the targeted event in the cartoon and if nothing else is available to be the target.

- a) Normally IRR but coded as target

(b) **For control items**

Control items elicit utterances expressing Manner and are therefore not coded in entirely the same way. Although the Richness criterion is still the first to be used, the Manner component becomes more important than the Path component as far as the primacy criterion is concerned. The three examples below show the various target utterance choices depending on semantic information provided.

- TM > M: TM is coded as CIR and M as MIR
He is sort of walking (M) [=MIR] and then he is walking across the screen (TM) [=CIR].
- TM > T: TM is coded as CIR and T as MIR
He is crossing the screen (T) [=MIR], he is running across (TM) [=CIR].
- M > T or MM > TT: M is coded as CIP (primacy of Manner), and T as MIP
He is running (M) [=CIP] and going across the screen (T) [=MIP].

If Manner has not been expressed by the participant initially but is produced following a specific question, the response can count as a target (CIQ), but the code LS in field 3 indicates that it is a response to a specific question and the issue is noted in a comment in field 16.

(c) **For all items**

There is an exception to the general principle concerning MI/CI pairs when two criteria are in conflict. In particular, the combination CIR + MIR + MIP is possible if the target is both richer than one and more relevant than another part of the response. In this case, the richness criterion is always given preference for the target (CIR).

- *He's climbing the hill with a rubber ring he's walking very slowly and he goes all the way up to the top pushing the rubber ring. He walks up the hill with a rubber ring [MIR] he walks very slowly [MIP] and he goes up to the top pushing the rubber ring [CIR].*

2.1.3 Field 3: Response type

For a given item, this field indicates:

- 1) If the coding line corresponds to a spontaneous response (LR) or to one answering a general (LG) or specific (LS) question.
- 2) The rank of the coding line when several are available for a given response, beginning with 01. In this case, the numbering of coding lines is consecutive for a given item, whatever type the response may be. If there is only one coding line, this is indicated by 01. Numbering takes account of all coding lines corresponding to an item, excluding lines corresponding to irrelevant content (IRR) and those corresponding to the absence of a component (CIB0/NO, etc.).

It should be noted that utterances coded as LS will be excluded from the analysis, and only spontaneous utterances or responses in answer to general questions will be included. Therefore, the LS coding should be used exclusively with questions that clearly ask for specific information.

LRn Coding lines corresponding to a spontaneous response or to an utterance following a presentative question, or a question to start the response, such as 'What happened?', 'Qu'est-ce qui s'est passé?' 'Was ist passiert?'

- *EXP: Okay, so tell me what happened?
- *SUF: A little mouse tiptoed across the screen. (=LR01)
- *EXP: Und das Eichhörnchen?
- *SUF: Das hüpfte da hoch. (=LR0x)
- *EXP: Et alors?
- *SUF: La valise elle a glissé. (=LR0x)

LGn Line corresponding to a response in answer to a general question: And then? Und dann? And after

that? *Et après? Und danach?* And at the end? *Et à la fin?*

LSn Line corresponding to a response in answer to a specific question:

- *EXP: *How did the ball go across?*
- *EXP: *Comment il a traversé?*
- *SUI: *das baby ist geklettert.* (=LR01)
*EXP: *geklettert?*
- *SUI: *nein, gekrabbelt.* (=LS02)
- *SUI: *Popi est monté sur la dune.* (=LR01)
*EXP: *Et comment il a fait pour monter?*
*SUI: *Il a marché.* (=LS02)

LG01, LS02, LX03, etc.:

Responses to questions in a sequence comprising several coding lines.

Various potential problems

1. Differences between tasks and / or items

The criteria defining specific questions-responses can be different in part, depending on the task. For example, the same type of question could be specific in Hoppy (where several entities are involved) but general in Voluntary Motion (involving a single agent only). What is more, the criteria differentiating general and specific questions can be used more flexibly with very young children for whom significant support (scaffolding) is sometimes required in order to obtain output.

Examples:

- *And then what did he do?* (QG in Voluntary Motion events, but QS in Hoppy as likely to elicit further sub-events).
- *And then what did the ball do?* (QS in Hoppy as likely to elicit a further sub-event).
- *And then what did the squirrel do?* (QG in Voluntary Motion event).
- *Und wie macht er das genau? 'And how did he do that exactly?'* (QS in all tasks).
- To elicit a motion event, the experimenter asks the child: *Sie ist weggeflogen? 'Has she flown away?'* when the squirrel climbs down the tree. (QS).
- *Was macht das? What is it doing?* (in control QG items).

2. Scope of questions

Problems can occur in relation to the scope of questions. For example, specific questions may be followed by responses concerned with something else (In these cases, code QG).

2.1.4 Field 4: Situation type

For a given item, this field indicates the situation type presented.

SS The situation is static (*he's on the road, elle est en vélo, Popi ist im Wald*).

SD The situation is dynamic (*she's cycling across the tracks, il court sur la route, er bleibt stehen*).

SX The situation is indeterminate or impossible to classify as belonging to the first two categories (for example in cases of the total absence of a verb which cannot be inferred from the context).

Please note that for this coding negation is to be ignored (it is indicated elsewhere, see field 5).

- *The basket didn't fall* coded as: *The basket fell* (=SD).
- *Il ne marche pas, il roule* coded as: *Il marche* (SD), *il roule* (SD).
- *Am anfang ist er da oben, aber diesmal nicht an der linken Seite*
coded: *Am anfang ist er da oben* (SS), *aber diesmal ist er an der linken Seite* (SS).

2.1.5 Field 5: Verb

This field is used to indicate the verb type, the semantic information expressed (after #+ and separated by + signs), the infinitive of the verb (after the | bar), and other information relating to the context of the predicate in

which this verb is inserted (after §- and before the colon). Some languages such as Chinese and Korean use high frequencies of serial verb constructions in the expression of motion events. For these languages, it is not always possible to decide on the status of a linguistic item as being verbs or satellites, a distinction of importance in our coding. To see how we dealt with these ambiguities, please refer to volume two of the manual which includes an appendix on the coding of verb-like parts in these languages.

To help the coder decide on the verb type, examples from frequently used verbs in the three main languages with their coding have been included in Appendix x in Volume 2.

NB: Verb coding may have consequences for Figures and Grounds (see fields 14 and 15).

The general format of the field is as follows:

type#+info+info+info...|infinitive§-additional information:

- *The squirrel hopped to the tree* VMM#+Mmv|hop
- *Il monte la bouée* VMP#+Cse+Tve|monter§-trans
- *Popi steht in einer Stadt* VSS#+Mpo|stehen

Verb type

The verb type is coded by an abbreviation of three capital letters, beginning with the letter V at each time:

VSS#+...|infinitive Static verb

(e.g., there be, to have, to sit-static, *être, se trouver, stehen, sitzen, da sein*).

VMM#+...|infinitive Dynamic motion verb, voluntary or involuntary, but other than VCM or VMP

(e.g., *run, hop, roll*-intransitive, *sit-down, courir, partir, monter*-intransitive, *tomber, laufen, fallen, radfahren, aufstehen*)

Note: He climbs up and 0 back down again. In cases of simple anaphora, the verb is coded VMM both times.

VMP#+...|infinitive Caused motion verb

The verb or verbal construction expressing a motion event caused by an explicit external force without an explicit causative construction (e.g., *roll*-transitive, *push, pull, drag, rouler*-transitive, *pousser, tirer, mettre, schieben, ziehen, rein tun*).

VCM, VCP, VCX#+...|make-infinitive Explicit causative constructions (*make roll, faire rouler*)

The first two letters (VC) indicate that this concerns causative constructions.

The third letter corresponds to the type of subordinated verb in the construction (VCM = motion verb, VCP = path verb, VCX = other dynamic verb). After the vertical bar, the causative verb is indicated (*make, faire, machen, lassen* etc.), then the infinitive of the subordinated verb, separated by a hyphen.

- *he makes the ball roll* = VCM#...|make-roll
- *il fait manger le chien* = VCX#...|faire-manger
- *il lui fait prendre le chien* = VCP#...|faire-prendre

VXX#+...|infinitive Other dynamic verbs

All other dynamic verbs not (directly) involved with motion

(e.g., *eat, grab, lick, hold, manger, toucher, dormir, s'arrêter, apparaître, disparaître, zerbrechen, stehen bleiben, treffen*).

VSZ, VMZ, VPZ, VCZ, VXZ#+...|infinitive Absent and inferred verb

An absent verb whose infinitive is inferred by the linguistic and/or sometimes non-linguistic context, for example, following a question: [*Where is he going?*] *Into his house*. Verb = VMZ...|go

- The third letter (Z for 'zero') indicates the absence of the verb.
- The second letter indicates the type of verb when this can be inferred (VSZ = static, VMZ = voluntary

- motion, VPZ = caused motion, VCZ = causative construction, VXX = other dynamic verb).
- Do not use parentheses around absent and inferred verbs on the coding line.
- Do not code Z in the case of simple anaphora (*the tyre goes back into the cave and Hoppy too*). (cf. also earlier remark regarding anaphora).
- Include cases of ellipsis in German of the type *Er möchte rüber*, (Meaning: *He wanted to go across*)
- Also use this code if transcription problems (xxx) have been noted in the file.
- Code the semantic content of the inferred verb, as exemplified below:
- *A present, a sand dune.* VSZ#|be§-pres (for: *there is*)
- *Die Raupe ist auf den Stengel* (Accusative case). VMZ#|gehen§-sein
- *Popi hat das Geschenk zum Hügel hoch.* VPZ#+Cse|haben§-tun
- Use code FZ for the inflection (see field 6).

VZZ#|indet Absent and indeterminate verb, that cannot be inferred.
Put *indet* after the bar and use FZ for the inflection (see field 6).

Specific cases of complex constructions

Some complex verbal expressions are not segmented, whatever the language and the subordinate verb, and they are coded as a whole. These cases fall into three particular types (see also field 6):

a) **Causative constructions:** of the type *make+INF* or *faire+INF*: *make slide*, *faire rouler*, etc.

In this case, the two verbs are indicated after the bar and the inflection coded is that of the verb *make/faire*:

- *He makes the trunk slide down the hill* = VCM...|make-slide...:PR
- *Il a fait glisser la valise* = VCM...|faire-glisser...:PC

b) **Constructions comprising a modal or more generally a psychological verb:**

Examples of verbs in this category are: *be able to*, *need to*, *want to*, *decide*, *penser*, *savoir*, *wollen*, *wissen*, etc.

In this case, the infinitive of the coded verb is indicated after the bar and the main verb is indicated after the §-sign followed by its inflection (see also Field 6 for the coding of the inflection, below). For example:

- *He wants to swim* = VMM...|swim§-want...:PR
- *Il pouvait nager* = VMM...|nager§-pouvoir...:IMP
- *Er wollte essen* = VXX...|essen§-mögen...:PTT

c) **Complex constructions indicating a temporal-aspectual value:**

Examples of verbs in this category are: *start to*, *continue to*, *continuer à*, *se mettre à*, *finir de*, *anfangen*, *dabei sein*, etc. These cases are normally coded similar to constructions including modal or psychological verbs as in (b) above:

- *He starts to walk* = VMM...|walk§-start-to...:PR
- *Il a continué à courir* = VMM...|courir§-continuer-à...:PC

Nevertheless, some of these constructions, such as those including *continue* in the below examples, can be used to imply a context of ellipsis or anaphora, in which no subordinate infinitive verb is used. In these cases, the verb *continue* is indicated after the bar, and this verb is coded as being of the VXX type. This leads to the following possible codings of structures of this type:

- If the subordinate infinitive verb can be inferred, it is indicated in parenthesis with the §- sign as in:
He runs... and he continues = VXX#|continue§-(run)
- If the verb *continue* is used in a very vague way and if the infinitive verb cannot be inferred, the annotation *indet* is used after the §- sign to indicate that it was not possible to determine this verb:
Il continue jusqu'à la fin. = VXX#|continuer§-indet
- If the verb *continue* is used in relation with other content that does not include another verb, the content of the utterance is indicated, for example, when it concerns a nominal expression:

Il continue son chemin = VXX#|continuer§-son-chemin

Verb semantics: information type expressed

If using the coder menu in CLAN, this information is to be inserted manually once the whole line has been coded, except in the case of VZZ, for which the menu offers VZZ#| (nothing else is necessary). Enter the semantic information in the order below (i.e., always Cause first, then Path, then Manner).

Format:

In the event of a combination of information, add the various components, beginning each one with + ...#+info+info+info...|...:

-	<i>He's going to school:</i>	VMM# go:
-	<i>He climbs up the tree:</i>	VMM#+Mmv climb:
-	<i>He pulls the trolley into the cave:</i>	VMP#+Cse+Mac pull:
-	<i>Il monte à l'arbre:</i>	VMM#+Tve monter§-intrans:
-	<i>Il fait glisser le traîneau sur la rivière:</i>	VCM#+Cse+Mmp faire-glisser:
-	<i>Das geschenk schiebt der Popi aufs dach:</i>	VMP#+Cse+Mac schieben:
-	<i>Der junge ist über den Fluss geschlittert:</i>	VMM#+Mmv schlittern:

Relevant information likely to be expressed by the verb

Relevant to Cause:

Cse Cause of motion, for example, expressed in verbs such as 'put' or in causative constructions such as *faire rouler*. Cause is also coded in verbs expressing the Manner of the cause, such as *push* or *pull* (which are doubly coded for Cause, and Manner of the cause (or Agent action), see the *Mac* code below).

Relevant to Path:

Tve Vertical Path (*descend, monter, descendre, tomber, erklettern, steigen*).
Tbo Path with boundaries (*cross, leave, arrive, traverser, partir, arriver, sortir, entrer, apparaître, disparaître, überqueren, verschwinden, sich schleichen*).
Tdi Directional Path (*avancer*).
Tdx Path with deixis (*come, bring, venir, kommen, bringen*).
Tso Source Path (*s'envoler, démarrer, sich entfernen*).
Tgo Goal Path (*land, approach, approcher l'arbre, landen*).
Tpa Other aspects of the Path (*to make one's way, passer*).

Relevant to Manner:

Mpo Manner with static position (*sit, être assis, stehen*=VSS#+Mpo) and change of position (*sit-down, s'asseoir, sich hinlegen* =VMM#+Mpo). Includes the four standard positions (*sitting, standing, lying, hanging*) and other relevant cases (*to be horizontal, to be vertical*).
Mmv Manner of motion (*run, skip, cycle, courir, marcher, nager, skifahren, rollen-intransitive, spazieren*).
Mma Manner of the Agent – Hoppy series only (*walk*).
Mmp Manner of the Object – Hoppy series only, for example:

-	<i>roll, slide, rouler (intransitive) - The table slides, la balle roule:</i>	Mmp
-	<i>roll, rouler (transitive) - Hoppy rolls the ball, rouler la balle:</i>	Cse+Mmp

Relevant to the Manner of an action by an external force leading to the displacement of another entity:

Mac Manner of the agent action leading up to the motion – if the Manner of the cause is expressed, it necessarily leads to a double coding Cse+Mac, for example *push, drag, shove, pousser, tirer, (an)schubsen, kippen, kicken*.

Others:

Oth Coding problems (rare cases).

#| None of these components is expressed (to go, *aller*, *gehen* – with the exception of cases where *gehen* is used without a particle to indicate a departure: *und dann geht er*, *and then he goes*).

Only code information conveyed by the verb; more information can be expressed elsewhere (satellites, Figure and Ground). For example, ‘The horse took the last fence very cleanly’ probably implies *leap over*, but this is not expressed by the verb *to take*. Note also that some information can seem to come from the utterance but is actually the product of inferences by the coder (for example ‘walk in the cave’ *aller dans la grotte* understood as ‘enter the cave’).

There can be redundancy. The information above can also be expressed by other elements of the utterance, such as expressions clarifying Manner or an instrument serving to bring about motion. In these cases, each occurrence should be coded with the appropriate code concerned (see also fields 8 to 12). Examples:

- *to pull with a string:* VMP#+Cse+Mac|pull:... EO1/1&+Mac|with-string(inst)
- *marcher à quatre pattes:* VMM#+Mmv|marcher:... :EO1/1&+Mmv|à+quatre+pattes

In the case of causative constructions with the verb ‘to make’ *faire*, take account of the information expressed by the second verb (e.g.: ‘make roll’: VCM#+Cse+Mmp|make-roll).

Distinctions indicated after the infinitives of verbs used (after the §- sign)

FORMAT: ...|infinitive§-distinctions:...

After the bar, the infinitive of the verb (or the regular quotation form for the language coded) is indicated, then the §- sign, then other information concerning the predicate (see also below), which can be of different types, cf. Table 2 below.

• Particles and adverbials

Indicate spatial particles after the §- sign, but do not indicate other satellites. Note that semantic information expressed in particles and adverbials is coded on the relevant fields (8-12). For example:

- ‘to run away’ ... VMM#+Mmv|run§-away ... EL1/1&+Tbo|away-z
- ‘to swim across’ ... VMM#+Mmv|swim§-across ... EL1/1&+Tbo|across-z
- ‘to jump up from the bottom to the top of the tree’
... VMM#+Mmv|jump§- up...EL1/3&+Tve|up-z... EP2/3&+Tso+Loc|from-bottom-z...
EP3/3&+Tgo+Loc|to-top-of+NP...

• Temporal-aspectual marks

Indicate modals and temporal-aspectual markers.

He wants to go up... |go§-up-want

Il commence à descendre... |descendre§-commencer-à

• Frequent and polysemous verbs

For polysemic verb roots that can be used in different contexts (motion events or non motion events), after the infinitive, indicate the relevant information that justifies the verb type coding:

- ...*to get to the nest* ...VMM#|get§-to-nest
- ...*faire une promenade* ...VMM#+Mmv|faire§-une-promenade

NB: Verbal constructions with *faire* are not coded in the same way, depending on whether they are causative (VCM#+Cse+Mmp|faire-rouler) or not (Voluntary motion): VMM#+Mmv|faire§-une-promenade.

With the following verbs, certain additional codes may be used to exhaustively identify their use (none of these verbs should occur without annotation after the §- sign).

be (static situation SS and verb VSS):

Localisation: -loc (*Hoppy is on the left hand side of the road, the cheese is on the table, il est là, il est sur la table, der Schubkarren ist hinter Popi, die Katze ist auf dem Mast.*)

Possession: -poss ('It's his', *C'est à moi = C'est le mien, das ist ihres*). In the French case, do not code the preposition *à*. In certain cases of set expressions (*C'est à moi = 'C'est mon tour'*), code *-poss-formule* after the § sign and indicate {*formule*} code in the comments (field 16).

Presentative: -pres (*There is a caterpillar, C'est Popi qui monte la dune, dabei ist da irgendwie ein Sandhügel gewesen*. Includes cleft sentences).

Mixed cases: cases containing more than one annotation.

- *There was some honey in there* (pres-loc)
- *C'est la bouée de Popi* (pres-poss)
- *Et là c'est la bouée de Popi* (pres-poss-loc)
- *Da/Hier ist ein Mann* (pres-loc)
- *Da ist noch ein Sack und da noch ein Sack* (pres-loc).

Verb 'to be' used for 'to go' (This case only exists for past tense inflections).

- *He's been at the top of the tree and then came back down:* SD:VMM#|be§-go
- *Il a été en haut de la colline:* SD:VMM#|être§-aller.
- *Der ist hoch.* ('He has gone up.'): SD:VMM#|sein§-gehen.

Other cases: specify what follows.

- *He is small:* VSS#|be§-small (be ill, handsome, inclined, etc.)
- *Il est debout:* VSS#+Mpo|être§-debout
- *Il est à quatre pattes:* VSS#+Mpo|être§-à-quatre-pattes
- *Die Schranken sind offen:* VSS#|sein§-offen
- *... weil das wieder so rutschig ist.* VSS#|sein§-rutschig

Uncertain: -indet

Have (situation SS and verb VSS):

Possession: -poss (*Hoppy has a shopping trolley, Popi a un cadeau, das Mädchen hat Schlittschuhe*).

Presentative: -pres for existential constructions **y-avoir** (*il y avait un chat*). With these constructions, do not code *y* as a spatial expression (see fields 8-12).

Other cases: specify what follows (*avoir envie, avoir mal, avoir peur*).

Combination -loc-poss

- *Il a une table devant lui.* #VSS/avoir§-loc-poss
- *Er hat einen Stuhl hinter sich.* #VSS/haben§-loc-poss

Uncertain: §-indet

make (SD situation):

Verbal anaphora: if verbs like *do* and *make, faire, tun, machen*, replace another verb, indicate this other verb in parentheses, and code its category.

- *He wants to grab the babybirds and climbs up to do so.*
Il le met là, et il fait vite.

Other cases: specify what follows (*make a cake, faire des galipettes > VXX*).

Uncertain: -indet

go, gehen, aller (SD situation):

The verbs *go, gehen, aller* are considered in this coding manual as being 'neutral' in the sense that they only express motion itself despite certain exceptions below (and despite the fact that other approaches consider these verbs as intrinsically deictic). The German verb *gehen* can have several senses though, including as a neutral verb where it only expresses motion. Most of the time, and when there is uncertainty, this verb is therefore coded as a verb of motion without semantic content: VMM#|gehen.

- *dann geht er rüber* 'then he goes across': VMM#|gehen

One exception concerns the use of *gehen* to express a departure (equivalent of *weggehen*). In this context, and if the utterance does not contain a verb particle, this verb is coded as expressing information relating to the path (particularly *Tbo*), with the context being indicated by the annotation *-exit* after the § sign.

- *und dann geht er* 'and then he goes': VMM#+Tbo|gehen§-exit

• Other distinctive verbs

In the case of caused motion events, some verbs seem to imply both the Manner of the agent's action and that of the object, particularly in cases with the verbs *to drag*, *traîner* (transitive)

- *to drag the bag*, *traîner le sac*: VMP#+Cse+Mac+Mmp
- *faire traîner le sac* = VCP#+Cse+Mmp
- *le sac traîne* (intransitive) = VMM#+Mmp
- *to walk the dog*, *promener le cheval* = VMP#+Cse+Mmp|promener

In certain cases of voluntary motion events, the potential indication of a Path is not coded, particularly when it is not relevant for the motion events targeted. For example, verbs such as *se retourner*, *se tourner* or *tourner* are coded without semantic content when they refer to a change in position (*se retourner* for 'to go back down'): VMM#|se-retourner, VMM#|se-tourner, VMM#|tourner. Similarly, the below examples are coded VXX:

- *to hold*, *tenir*, *halten*

Whereas the following set of verbs are considered to have some relevant semantic information:

- *to bring*, *emmener*, *ramener*, *apporter*, *emporter*, *bringen*: coded VMP#+Cse+Tdx|
- *to carry*, *tragen*: coded VMP#+Cse|

• Special coding with regard to infinitives and verbal constructions in German

a) In cases where objects (*Eis*, *Fahrrad*) are incorporated (*eislaufen*, *fahrradfahren*, etc.), the full verb is coded after the bar and the verb root is noted after the §- sign.

- *Er ist eisgelaufen.* = VMM#+Mmv|eislaufen§-laufen
- *Er ist snowboardgefahren.* = VMM#+Mmv|snowboardfahren§-fahren

b) When a verb construction implies a full verb accompanied by a neutral verb, in particular *gehen* 'to go' or *tun* 'to make', the full verb is indicated after the bar and the neutral verb after the § sign.

- *Er tut baden.* = VMM#+Mmv|baden§-tun
- *Er geht schwimmen.* = VMM#+Mmv|schwimmen§-gehen

c) In German, numerous verbs of motion are accompanied by the prefix *zu-* and a prepositional expression. In these cases, the verbal construction and its complements express a motion carried out in a specific direction (indicated in the prepositional expression). This type of construction is noted after the §- sign but the semantic content concerning the Path (Tdi) is coded in the satellites separate from the verb (Fields 8 to 12).

- *Er krabbelt auf den Baum zu.* = VMM#+Mmv|krabbeln§-zukrabbeln-auf+NP

d) The following coding is used for constructions of the kind: *an* + past participle (motion verb) + *kommen*. The construction is indicated after the bar, and the two infinitives are noted after the §- sign; firstly, the verb *kommen* (Tdx), then the prefixed verb (most of the time indicating the Manner of the motion):

- *Er kommt angefahren.* = VMM#+Tdx+Mmv|angefahren-kommen§-kommen+fahren:..... {formule}

• 'Non spatial' values

Certain verbs which usually have a spatial meaning can also have another sense, for example:

- VMM verbs: *walk* (*let me walk you through the contents of the book*); *run* (*the motor is running smoothly again*); *arriver* (*j'arrive pas à ouvrir*); *aller* (*ça va bien ?*); *marcher* (*la radio marche plus*); *gehen* (*und wie geht's dir?*)
- VMP verbs: *put* (*I put it to you that..*); *mettre* (*je mets la musique*)
- VSS verbs: *y-être* (*ça y est! C'est fini*).

In these cases, leave the normally expected code (VMM, VMP, VSS), but indicate the ‘non-spatial value’ code -*vns* after the §- sign, as well as after the bar for coding the Figure and the Ground (see fields 14 and 15). Also add the *{formule}* code in the commentary (see field 16).

- **The role of context**

Certain verbs such as *catch*, *turn*, and *voler* in the examples below can be coded VMM, VXX or VMP according to their sense in the context.

-	<i>catch an object that is falling</i>	=VXX
-	<i>catch someone who is running</i>	=VMM
-	<i>turn left</i>	=VMM
-	<i>turn the page</i>	=VMP
-	<i>voler dans les airs</i>	=VMM
-	<i>voler</i> (‘steal’)	=VXX

In French, ambiguities between VMM and VMP can be linked to the use of the auxiliary with a clitic pronoun: *a roulé / il l’a roulé*, or *il a monté / il l’a monté*: It is better to use VMM if the context does not allow disambiguation between VMM and VMP, and to note the following code in the brackets at the end of the line {ambigv-vmm/vmp}.

- **Ambiguity linked to the immediate future tense (French)**

Certain cases of *aller* are ambiguous between the immediate future tense and a motion verb (*il va nager = il va bientôt nager* or *il se déplace à la piscine pour nager*). Choose the most likely interpretation in the context and indicate the ambiguity in a comment at the end of the line: {ambigvm} (see field 16).

- **Reflexive verbs in certain causative constructions indicating the Manner of the motion**

Certain causative constructions are interpreted as concerning Manner of voluntary motion, rather than caused motion. This is the case in constructions containing a reflexive verb such as *to let oneself +INF*, *se laisser + INF*, *se faire + INF*, or *sich lassen +INF*. The coding indicates these cases by the code ‘causative-refl’ after the §- sign. Example:

<i>He let’s himself slide down the riverbank</i>	VMM#+Mmv... let-oneself-slide§-causative-refl
<i>Il se laisse glisser</i>	VMM#+Mmv... se-laisser-glisser§-causative-refl
<i>Er lässt sich runterfallen</i>	VMM#+Mmv... sich-fallen-lassen§-causative-refl

Likewise, causative constructions or caused motion verbs expressing a displacement forced by part of the agent's body are coded as concerning Manner of voluntary motion. The coding indicates these cases by the code ‘causative-body’ after the §- sign. Example:

<i>A girl skates forward by moving her right leg</i>	VMM+Mmv... move§-leg-causative-body
<i>Une chenille avance en levant son dos</i>	VMM+Mmv... lever§-dos-causative-body

- **Negation**

Negation is ignored in verb coding (as in situation type, see field 4), but it is indicated after the infinitive of the verb by: §-*neg*. Examples:

- *The basket didn’t move or fall.*
Infinitives = *move§-neg* and *fall§-neg*
- *Il ne marche pas, il roule.*
Infinitives = *marcher§-neg* and *rouler*
- *Die Raupe klettert nicht sondern schleicht.*
Infinitives = *klettern§-neg* and *schleichen*

- **Static use of dynamic verbs (*fictive motion*)**

Certain dynamic verbs can be employed in a static sense, particularly in relation to spatial Grounds. Code SD for dynamic situations, add *§-sta* after the infinitive of the verb and note {fictive-motion} in the brackets at the end of the line.

- *The road runs along the old wall* VMM#|run§-along-sta... : {fictive-motion}
- *Une route va vers l'arrière-plan* VMM#|aller§-sta : {fictive-motion}
- *Une rue descend la colline* VMM#+Tve|descendre§-sta : {fictive-motion}

- **Transitive and intransitive verbs**

For certain motion verbs, note after the paragraph sign (§-) whether they are employed as a causative verb (§-**trans**) or not (§-**intrans**) – particularly for those that can be used in both ways. The codes ‘trans’ and ‘intrans’ refer to a notion of ‘semantic transitivity’ (and are not purely grammatical) and should be used directly in connection with the codes VMP and VMM (verb types), with which they are partially redundant, serving to aid analyses that concern a single sub-field. Among verbs that allow these two uses (for example *roll*, *monter*), those that are coded VMP are always ‘trans’ and those that are coded VMM are always ‘intrans’:

- *He rolls the ball down the hill* (causative transitive): \$:...VMP....§-trans
- *The ball rolls down the hill* (non causative intransitive): \$:...VMM....§-intrans
- *Il monte la valise* (causative transitive): \$:...VMP....§-trans
- *Il monte la colline* (non causative intransitive): \$:...VMM....§-intrans
- *Il monte à l'arbre* (non causative intransitive): \$:...VMM...§-intrans
- *Er rollt das Rad ins Haus* \$:...:VMP#+Cse+Mmp|rollen§**trans**:...
- *Weil es rollt* \$:...:VMM#+Mmp|rollen§-**intrans**:...

- **Idiosyncratic uses related to Field 5 - some examples**

§-idiov:

Idiosyncratic verbs: Indicate *-idiov* after the §- sign, and the verb that should have been used between hyphens, if this can help to interpret the coding. You can also mark idiosyncrasies with {*idiov*} in the comments (field 16). Examples:

Formal idiosyncrasy:

- *Er tut die Schubkarre runter fahren. He rolled the wheelbarrow down.* = tun§-rollen-idiov + {tun-cons}
- *Il sauve* (for *Il se sauve*) = VMM#+Tbo+Mmv|sauver§-idiov-se-sauver:

Functional idiosyncrasy:

Some verbs, although correct in form, can be employed in an idiosyncratic way, for example the dynamic use of a normally static verb: *j'étais (=j'allais) me coucher: être§-idiov-aller*

Prefixation

By definition, a prefix is a morpheme appearing at the front of a lexical unit (the verb root in this instance) which gives it certain semantic properties. Depending on the language, prefixes can be more or less frequent and especially more or less productive and may or may not be treated as elements having the status of real satellites. In some languages (for example, in Polish, Russian or Greek), they form a productive paradigm in which consideration can be given to whether they are satellites on the same level as a particle (of the type *up* or *away* in English). Conversely, in other languages, they have little effect and relate to a residue inherited from diachrony, and only partially function as satellites (for example, *ac-courir* comes from a prefixation in French), or have lost all their productivity (for example, *re-cevoir* is synchronically no longer considered prefixation). In languages where various prefixation types co-exist, coding differentiates between different cases in order to code prefixed verbs according to how they break down morphologically. These codes appear after the paragraph sign. Examples of such prefixes are provided in Volume 2 section x. Where prefixation occurs, this can lead to

a change of the meaning of the verb root. In some languages, it can be important to code for this phenomenon separately from the previously discussed codes for the verb field. Examples on how to code these effects are again provided in Volume 2 under section xxx.

Table 2 below provides a summary of the more complex verb-coding rules as described above.

Table 2: Overview of complex verb-coding rules

Type	Example	Code
Semantic transitivity information: Used for motion verbs with multiple meanings (VMM / VMP)		
'Intransitive' verb	<i>The ball rolls (=VMM)</i>	roll§-intrans
Transitive causative verb	<i>He rolls the ball (=VMP)</i>	roll§-trans
Causative construction		
Simple causative construction	<i>He makes the ball roll</i>	VCM... make-roll
Causative construction with reflexive treated as VMM	<i>He lets himself slide</i>	VMM... let-oneself-slide§-reflexive
Subordinate verb to a psychological verb		
	<i>He wants to swim</i>	swim§-want
	<i>She can run</i>	run§-can
Other compound verbal expressions		
	<i>Sie läuft Schlittschuh (to skate)</i>	schlittschuhlaufen§-laufen
	<i>Sie fährt Ski (to go skiing)</i>	skifahren§-fahren
	<i>Der Junge kommt gerannt (to arrive)</i>	gerannt-kommen§-rennen+kommen
	<i>Der geht tauchen (to go diving)</i>	tauchen§-gehen
	<i>Er kommt am Baum an (to arrive)</i>	ankommen§ -an
	<i>Der tut radfahren (to go cycling)</i>	radfahren§-fahren-tun
	<i>Er geht auf den Baum zu (to go near)</i>	zugehen§-zu
	<i>Sie hat Schlittschuhe an (to wear skates)</i>	anhaben§-schlittschuhe
	<i>Il continue à marcher</i>	marcher§-continuer-à
	<i>Elle fait du ski</i>	faire§-du-ski
Verb+other linguistic means		
	<i>He stands still (adverb)</i>	stand§-still
	<i>He runs away</i>	run§-away
	<i>He is running around</i>	run§-around
	<i>Puis il continue sa route (verb object)</i>	continuer§-sa-route
	<i>Die Flasche geht kaputt (to break)</i>	gehen§-kaputt
Negation		
	<i>Elle roule pas, elle tourne</i>	rouler§-neg
	<i>Er unternimmt nichts (he undertook nothing)</i>	unternehmen§-neg(nichts)
The nature of indeterminate verbs		
	<i>Et puis il fait</i>	faire§-indet
'Non-spatial' value		
	<i>The car runs smoothly</i>	run§-smoothly-vns
Idiosyncratic use		
	<i>il faut que je m'en bouge</i>	s'en-bouger§-idiov
Other information for specific verbs (être, avoir, faire)		
	<i>He's on the hill</i>	be§-loc
	<i>There is a cat</i>	be§-pres
	<i>Il y a un chat</i>	avoir§-pres
	<i>C'est à moi (= 'it's mine')</i>	être§-poss
	<i>He has a suitcase</i>	have§-poss
	<i>There's a little man on a hill</i>	be§-pres-loc

2.1.6 Field 6: Verb Inflection

A group of codes serve to indicate the verbal inflections carrying temporal-aspectual information in accordance with a specific list for each language. The lists presented in the tables below for English, French and German are not exhaustive and can be added to if necessary. For proposed coding of verb inflections in other languages, please refer to Volume 2, section x. Other more general codes are presented further on.

INFLECTION BY LANGUAGE

English temporal-aspectual verbal inflection

PR	Simple present	<i>he runs</i>
PRG	Present progressive	<i>he is running</i>
PAS	Simple past	<i>he ran</i>
PAG	Past progressive	<i>he was running</i>
PRF	Present perfect simple	<i>he has run</i>
PRFG	Present perfect progressive	<i>he has been running</i>
PPA	Past participle	<i>a bird located on the fence</i>
PPF	Past perfect simple	<i>he had run</i>
PPFG	Past perfect progressive	<i>he had been running</i>
FUT	Future simple	<i>he will run</i>
FUT2	Future simple 2	<i>he will have run</i>
FUTG2	Future progressive 2	<i>he will have been running</i>
FIPR	Future simple immediate	<i>he is going to</i>
FIPS	Future immediate past	<i>he was going to</i>
INPG	Infinitive progressive	<i>he seems to be running</i>
MOD	Modal auxiliary	<i>he can run</i>
CON	Conditional simple I	<i>he would run</i>
CONG	Conditional progressive I	<i>he would be running</i>
CON2	Conditional simple 2	<i>he would have run</i>
CONG2	Conditional progressive 2	<i>he would have been running</i>
PRA	Temporal-aspectual constructions in the present tense	<i>he begins to run, stops running</i>
PSA	Temporal-aspectual constructions in the past tense	<i>he began to run, stopped running</i>
PPR	Present participle	<i>he sits there watching them, he runs down dragging the bag, ...</i>
INF	Infinitive	<i>to run</i>
IPT	Imperative	<i>run Hoppy run!</i>

French temporal-aspectual verb inflection

Code	Inflection	Example
PR	Présent	<i>Il court</i>
PC	Passé Composé	<i>Il a couru</i>
IMP	Imparfait	<i>Il courait</i>
PS	Passé simple	<i>Il courut</i>
PQP	Plus que parfait	<i>Il avait couru</i>
FUT	Futur	<i>Il courira</i>
INF	Infinitif	<i>Il arrive pour courir</i>
INP	Infinitif Passé	<i>Il était content de l'avoir fait courir</i>
IPT	Impératif	<i>Attention Popi, cours !</i>
PP	Participe Passé sans auxiliaire	<i>Il arrive de l'autre côté tout mouillé</i>
GER	Gérondif ou Participe présent	<i>Il arrive en courant, courant...</i>
FIP	Future Immédiat au présent	<i>Il va courir</i>
FII	Future Immédiat au passé	<i>Il allait courir</i>
PRS	Présent du subjonctif	<i>Il faut qu'elle court/aille courir</i>
CND	Conditionnel	<i>Il aurait couru</i>
AMB	Ambiguity	<i>tomber/tombé; tomber [=?tombé] coded INF- AMB-PP</i> (if the context cannot be distinguished and when both possibilities are indicated in the transcription)
PRA	Présent + expression temporo-aspectuelle verbale	<i>Il est en train de courir</i>
IMA	Imparfait + expression temporo-aspectuelle verbale	<i>Il était en train de courir</i>
PCA	Passé Composé + expression temporo-aspectuelle verbale	<i>Il s'est mis à courir</i>
FUA	Future + expression temporo-aspectuelle verbale	<i>Il finira par courir</i>
PAI	Infinitif lié au passé immédiat	<i>Il vient de courir</i>
FIC	Infinitif lié au futur immédiat où le verbe <i>aller</i> est au passé composé	<i>Il est allé courir</i>

German temporal-aspectual verbal inflection

PR	Präsens	<i>Er läuft</i>
PTT	PräTeriTum	<i>Er lief</i>
PFK	PerFeKt	<i>Er ist gelaufen</i>
PQP	PlusQuamPerfekt	<i>Er war gelaufen</i>
FUT	FUTur	<i>Er wird laufen</i>
PPR	Partizip PRäsens (I)	<i>Zum Baum laufend sieht er die Vögelchen</i>
PAP	PARTizip Perfekt (II)	<i>Er kommt gelaufen</i>
KJT1	KonJunkTiv 1	<i>Er sagte er sei der Schnellste</i>
KJT2	KonJunkTiv 2	<i>Er dachte sie wäre nicht zuhause</i>
INF	INFinitiv	<i>Laufen</i>

Codes relevant for all languages

FZ	Absence of verb	<i>The hill, the gift</i>	VZZ# indet:FZ
FOT	Verb lacking morphology	<i>He go down</i>	VMM# go§-idionom:FOT
AMB	Ambiguous morphology	<i>Il [puse]</i>	VMP#+Cse+Mac pousser§-idiom:AMB

1) Absence of morphology: FZ vs. FOT (in all languages)

In the absence of a verb (coded VZ), the code FZ is used for the inflection.

Total absence of morphology: A code (idionom) is used when there is a verb but no morphology. Such cases are coded FOT in the field of verbal inflection:

- *He go* = VMM#|go§-idionom....:FOT:... (Absence of morphology)

Incorrect or ambiguous morphology:

- *He going* = VMM#|go§idionoaux:PRG:

- *Il [=? entrer/entré]* = VMM#+Tbo|entrer§-idiom...:AMB

In some cases where the inflection is lacking, incorrect, or ambiguous, the intended inflection can be inferred by the context. In these cases, this inferred form is coded. If not, AMB is used instead.

- *He is go* = VMM#|go§-idiom ... :AMB:...
Ambiguity between 'he is going' and 'he is gone'.

- *He is swim* = VMM#+Mmv|swim§-idiom...:PRG:...
Interpreted as 'he is swimming'.

- *He put it there* = VMP#Cse|put:AMB:...
= VMP#Cse|put:PAS:...

2) Temporal-aspectual periphrastic expressions

These expressions are indicated after the paragraph sign (§-) and coded with the corresponding inflection (PRA, PCA...) (see also cases of inferences, Field 5, Section A above): walk§-start-to; run§-continue-to; roll§-finish; courir§-commencer-à; courir§-venir-de; courir§-continuer-de; courir§-finir-de; krabbeln§-anfangen; schieben§-aufhören

Il s'est mis à courir = VMM#+Mmv|courir§-se-mettre-à:PC

2.1.7 Field 7: 'VOICE'

To facilitate searches, the first letter of this field is always 'D' for diathèse, or 'voice'.

DAC Active voice (Example: *He is sliding, she is cycling, he is rolling the ball, il court, il pousse le ballon, il se laisse glisser, er schiebt den Tisch, er zieht das Auto, er hopft*)

DPA Passive voice (Example: *The wheel is rolled by Hoppy, Le ballon est poussé par Popi, der Kinderwagen wird von Popi gezogen*)

DOT Other or indeterminate voice, including reflexive, semi-passive, etc. (E.g.: *the bottle is broken*).

DPG Get-passive in English (E.g.: *The bottle got smashed by the ball*).

DMI Medium voice as in Greek.

DBP Bei passive as in Chinese.

2.1.8 Fields 8 to 12: Satellites

The term ‘satellites’ is used loosely to include all forms other than the verb that express relevant spatial information. Five fields are provided. If the utterance includes just one satellite, code it in field 8 and code the absence of other satellites in fields 9 to 12 (see codes below). If the utterance includes more than one satellite, code them in fields 8 to 12 in the order in which they are produced. The order marking of the satellites is already provided in the coding menu. Once the line has been fully coded, add the total number of satellites produced for this line manually in each of the fields. Example:

- EP1/1 = 1st satellite of the prepositional type in a coding line containing just one satellite;
- EP1/4 = 1st satellite of the prepositional type in a coding line containing four satellites.

The other codes in this field indicate the form type used, the information expressed (after the & sign and separated by +), and the form used in a standardised format (after the bar |) (*jusqu’au=jusqu’à*, *au=à*, *des collines=de*). Example:

*SUI: Mr Hoppy is taking a chair, from the right-hand side of the picture, into the cave on the left-hand side [c].
%cod: \$:CIB0:LR01:SD:VMP#+Cse|take:PRG:DAC:EP1/3&+Tso+Loc|from-right-hand-side-of+NP:EP2/3&+Tbo|into+NP:EP3/3&+Loc|on-left-hand-side- z:EZ4/3&|:EZ5/3&|:PM0: Fig|chair:Grd|right-hand-side-of-picture+cave+left-hand-side(picture):{ }

FORMAT: E...n/n&+info+info+...|form (E.g.: *EL1/2&+Tve|up*)

Form types (Also see Appendices)

EP1/n&+...|form

Preposition or spatial prepositional expression

- *in, under, above, towards, from, etc.*
- *chez, dans, sous, sur, vers, par, en, de, à, etc.*
- *auf, über, von, in, etc.*

The coding indicates whether or not there is a following noun phrase expressing an explicit Ground.

With explicit Ground, the code should end with the preposition produced and the symbol +NP: *to+NP, up+NP, into+NP, au-dessus-de+NP, à-droite-de+NP, à-côté-de+NP, jusqu’à+NP, zu+NP, über+NP*, etc.

Note: Some expressions can indicate locations as well as Grounds, such as internal location nouns (*top of, side of, back of, autre côté*, etc...).

- *to run to the top of the tree*
VMM#+Mmv|run...EP1/1&+Tgo+Loc|to-top-of+NP...Grd|top-of-tree.
- *traverser jusqu’à l’autre côté de la rivière*
VMM#+Tbo|traverser...EP1/1&+Tgo+Loc|jusqu’à-l’autre-côté-de+NP...Grd|l’autre- côté-rivière

Without explicit Ground, the code should end with -z (hyphen z, for ‘zero’): *above, au-dessus-z, à-droite-z, vers-la-droite-z, à-côté-z, en-haut-z*, etc.

- *Il va à droite de la colline*: indicate *à-droite-de+NP* vs.
- *Il va à droite*: indicate *à-droite-z*
- *Il est à la droite de la maison*: indicate *à-la-droite-de+NP* vs.
- *Il est à ma droite*: indicate *à-ma-droite*

EL1/n&+...|form

Spatial particle: These forms only exist in some languages, for example, in English and in German, but not in others, i.e., French.

Among the particles, include elements that are not followed by an explicit Ground (*He climbs up, he runs across*). Those which are followed by an explicit Ground are coded as prepositions (*He runs into the cave, He is swimming in it, he runs across the street*).

In German, leave compound forms made from two particles together (*herauf, herunter*). Code *an* in the following constructions and indicate with which verb they occurred after the (§-) sign.

- *er kommt angelaufen:* = EL&+Tgo|an§-kommen... {formule}
- *er schleicht sich an:* = EL&+Tgo|an§-schleichen
- *er kommt an:* = EL&+Tbo|an§-kommen

EX1/n&+...|form

Other spatial expressions. This category must be reduced at most to expressions of type:

- *inside, outside, dessous, dedans, dessus, devant-z, dehors, oben, unten, rechts, links.*
- deictics (*here, there, là, ici, hier, da, dort*).
- pronominal spatial expressions (French *en, y, où*).
- adverbial expressions (*backward, forward, straight, loin*).

Some expressions can indicate location as well as Ground, such as deictic expressions (e.g., *là*). In this case, indicate the expression in the satellite field (EX) with the coding +Loc. Do not forget to code the deictic expression in the 'Ground' field too:

...*aller là* = EX1/1&+Loc|là ... Grd|là

EN1/n&+...|form

Relevant noun phrases for the expression of motion, for example expressing Manner (*a jogger, un nageur, une patineuse, die radfahrerin*), or the Manner and the Path (*le grimpeur*). Only code when occurring inside the clause (even in cases of anaphora).

EO1/n&+...|form

Relevant but non-spatial expression. These cases are for the most part linked to the expression of the Manner of motion, expressing, for example, the parts of the body used, or speed of motion (*slowly, à quatre pattes, wie ein Affe, seitlich*). Include prepositions such as *with* if they are accompanied by relevant information, for example, significant referents in the stimuli (*He climbs up with a ball, she goes down with the cheese*). Exclude cases irrelevant to the expression of motion (*merrily*).

EZ1/n&|form

Absence of a relevant expression. In this case, always code ...:EZ1/0&|:....

This code is used in all fields 8-12 if the utterance coded contains no spatial information outside of the verb, or in the remaining fields if the utterance contains fewer than five expressions, for example:

- *Hoppy pulls a bag.*
- %cod: \$:CIB0:LR01:SD:VMP#+Cse+Mac|pull§-bag:PR:DAC:EZ1/0&|:EZ2/0&|:EZ3/0&|:EZ4/0&|:EZ5/0&|:PM0:Fig|bag:Grd|imp(roof):{ }

Forms relevant for specific languages only

- *Languages containing case marking*

For languages containing case marking resulting in different spatial meanings (for example, German), prepositions used with a given case, i.e., the accusative and dative in German are coded for these cases (EA) and (ED) respectively. Coding of case marking can be adjusted for other languages as necessary. If the case in such languages cannot be determined (incomplete, adverbial object, etc.), EP is used. The format is: EA1/

n&+.../form and ED1/ n&+...|form for accusative and dative marking respectively. Case marking errors should be noted in the comments in field 16.

- *Die maus ist auf den tisch (Accusative) draufklettert*
= VMM#+Mmv|kletteren:...:EA1/2&+Tgo|auf+NP:...:Fig|maus:Grd|da+tisch
- *Da war es noch, im kinderwagen.*
= VSS#|sein\$-loc:...:ED2/2&+Loc|in+NP:...:Fig|PRO(baby):Grd|kinderwagen
- *das baby ist von hier auf die seite rüber gekrabbelt.*
= VMM#+Mmv|krabbeln:...:EP1/3&+Tso+Loc|von-hier:...:Fig|baby:Grd|hier+seite

Information expressed (see Appendices)

Enter the relevant semantic information in the order below: Cause first, then Path, then Manner. Insert it once the whole line has been entered automatically, by typing it in manually. In this case:

- insert the relevant semantic information after the satellite type and before the vertical bar;
- finish off by indicating the total number of satellites coded in the line after the slash.

Relevant to Cause:

Cse Cause in the context of a caused motion. For example, in expressions designating the instrument used to provoke a displacement.

- *Durch magische kräfte angezogen; beim ziehen*

Relevant to Path:

Tve Vertical Path (*up/down, vers le haut, hoch, rauf, runter*).

Tbo Path with boundaries (*across, away, into, onto, out/out of, rein, hinein, rüber, weg, los(gehen), raus, [sie kommt] ins Bild (accusative), [er läuft] aus dem Bild, fort*).

Tdx Path with deixis (*back, zurück, hin, her, er geht auf den Baum zu*).

Tso Source Path (*from the house, from the clearing in the woods, partir de l'arbre, von rechts, er zieht die Schubkarre vom Schneehügel, davon*).

Tgo Goal Path (*walk towards the barn, aller à l'arbre, jusqu'à l'autre bout, nach links, er geht auf das Dach accusative), bis zu einem Baum, ankommen*).

Tpa Other aspects of Path (*around, le long de, along, entlang, schief, rum -, vorbei -*).

Be careful not to misuse Tgo in French prepositions, which are often ambiguous in which case they are coded Loc, for example:

- *Il est d'un côté de la rue* = VSS... EP&+Loc
- *Il est reparti de l'autre côté de l'écran* = VMM... EP&+Loc
- *L'ours court de l'autre côté* = VMM... EP&+Loc
- *Il va jusque de l'autre côté* = VMM... EP&+Tgo+Loc

Relevant to the Manner of a displacement:

Mpo Manner expressing the position (E.g.: *on all fours, up on two legs, straight, de structure verticale, le bébé est à quatre pattes, der Ball kickt die Kegel um, der Würfelturm bricht zusammen*).

Mmv Manner of motion – for Voluntary motion stimuli only (*walk slowly, in freestyle, in small leaps, bizarrement, seitlich, mit merkwürdigen auf und abbewegungen*).

Mma Manner of the agent – Hoppy series only (*fast, slowly, Popi marche à quatre pattes*).

Mmp Manner of the object – Hoppy series only (*around, round and round, avec un déplacement horizontal*).

Mmi Inferred Manner of motion (*there is a runner, he is going across the road; il y a une fille avec des patins et elle va jusqu'à l'autre côté*). In non-target utterances announcing a Manner, for example, with a static verb (see also further information below).

Mmx Manner by which an object is fixed to another object (*caught, attached*).

Relevant to the Manner of the action provoking a displacement:

Mac Manner of the agent action leading up to the motion (*with string, by rope, avec une corde, am Henkel, an einer Schnur, mit Seil*). In many of these cases, the Manner of the Cause implies the use of an instrument, coded (inst) after the form used (see below). Unlike the verb (which requires double coding Cse+Mac), the Mac code can be used on its own in the case of instruments indicating the Manner of the Cause: *He pulls with a rope*.

EO1/1&+Mac|avec-corde(inst)

Other information expressed:

Loc General location expressed (*he runs on the grass, il monte à l'arbre, er ist da*).
(*So he pulled the car up the roof at the house; il marche sur la route, Popi rollt den Ball vor sich hin, Popi zieht den Rennwagen über den Hügel. Bei dem Hügel runterschieben*).

Oth Coding problems.

None of these components is expressed:

In this case, nothing is written between the & sign and the bar, for example:

- *elle marche joyusement* = EO&|joyusement

Some additional information regarding the coding of semantic types of information:

1) 'Satellite pairs'

a) Pairs expressing two separate locations: Tso and Tgo Pairs

In general, the Tso and Tgo pair of codes is used when two satellites are used together to represent two locations, one indicating the source, and the other the goal:

- *He goes from the left side of the road (Tso) to the right side (Tgo)*
- *du haut (Tso) jusqu'en bas (Tgo), d'un côté (Tso) à l'autre côté (Tgo)*

b) Pairs expressing direction: 'satpair'

One important exception concerns the case where two expressions are used to express direction, rather than a combination of source-goal locations. For example, the expressions *from left to right, de droite à gauche, von links nach rechts* indicate the direction, and not the source and goal locations. These cases are marked by the additional code 'satpair' placed after the §- sign, indicating that it is a pair of expressions that are coded as a single satellite, for example:

- *He goes from left to right* = EP...Tdi|from-left-to-right§-satpair
- *Il va de bas en haut* = EP...Tdi|de-bas-en-haut§-satpair
- *Er läuft von links nach rechts* = EP...Tdi|von-links-nach-rechts§-satpair

c) Special cases: 'sat1' and 'sat2'

In certain cases, the use of a pair of satellites falls well within two distinct components (for example, +Tso+Tdi or +Tso+Tgo) but presents coding difficulties. These cases are often seen, particularly in French, where the entire utterance, which is formed by two prepositional expressions, appears in part to indicate a directional sense, despite falling within the expression of the location from which the direction is inferred. The coding of these cases fits neither the elements to be coded separately as in a) above, nor the elements to be coded as a satpair: In this case, the codes 'sat1' and 'sat2' are placed after the §- sign for each of the two expressions. Example:

- *De la gauche vers la droite* = Tso§-sat1..... Tdi§-sat2
- *De la gauche jusqu'à la droite, de la gauche à la droite* = Tso§- sat1..... Tgo§-sat2

d) Other complex expressions

Some complex expressions are coded together, whether they are used singly or as an element in a pair of satellites, for example: *top left, bottom right, d'en bas à droite, von links oben*.

- *And a cave at the bottom right of the screen* = EP1/1&+Loc|at+bottom-right-of-NP
Grd|bottom-right-of-screen
- *Il arrive sur le trottoir de droite* = EP1/1&+Loc|sur+NP
Grd|trottoir-de droite
- *Il arrive sur le trottoir d'en face* = EP1/1&+Loc|sur+NP
Grd|trottoir-d'en face

Other examples for complex expressions:

- *Elle arrive à droite de l'arbre*
EP1/1&+Loc|à-droite-de+NP...Grd|arbre
- *Il va jusqu'en bas*
EP1/1&+Tgo+Loc|jusqu'en-bas-z...Grd|imp(indet)
- *Jusque dans la grotte*
EP1/1&+Tgo+Loc|jusque-dans+NP...Grd|grotte
- *Il part du haut de la colline*
EP1/1&+Tso+Loc|du-haut-de+NP...Grd|haut-colline

2) Vague Manner

If the Manner is expressed verbally but in a very vague way or with onomatopoeia and/or a co-verbal gesture, no semantic content is coded with regard to Manner, but (nv) is indicated after the expression:

- *He went back down like that.* = EO&|like-that(nv)
- *Il monte comme ça* = EO&|comme-ça(nv)
- *Et après, il fait pfft* = VXX#|faire§-pfft(nv)
- *Il est comme ça.* = VSS#|être§-pres...EO&|comme-ça(nv)
- *Der ist so runtergeklert* = EO&|so(nv)

The code (nv) is not used in the following cases, where the semantic content can be inferred. For example, when an expression such as 'like, comme, wie' introduces an entity indicating the Manner of the utterance:

- *He walks like a baby* = VMM#+Mmv|walk...EO1/1&+Mmv|like-baby
- *Il se déplace comme un singe* = VMM#|se-déplacer...EO1/1&+Mmv|comme-singe
- *Er läuft wie auf Eierschalen* = VMM#+Mmv|laufen...EO1/1&+Mmv|wie-auf-eierschalen

or when a reference to Manner refers anaphorically to something mentioned previously, for example with expressions such as *in the same way, de la même manière*, etc.:

- *et il repart de la même manière:* = EO1/1&+Mmv|de-la-même-manière

3) Subordinate clauses treated as adverbial expressions

In certain cases, a complex utterance is treated as if the subordinate clause was a simple adverbial expression (coded EO) equivalent to "wie eine Raupe":

- *Sie bewegt sich, wie eine Raupe sich eben bewegt.* = EO&+Mmv|wie-eine-raupe

4) Inferred Manner (Mmi) and Attachment Manner (Mmx-spatial configuration)

The code *Mmi* indicates information that ‘announces’ a Manner of motion or caused motion expressed in an utterance that is not part of the target. The *Mmx* code indicates information that announces a spatial configuration and/or a Manner of attachment between two entities. These cases are often found in settings that include static situations coded in the BEG lines. In these special cases of inferred Manner or attachment Manner, indicate the Mmi or Mmx codes after the §- sign. Examples:

- *She is sitting on a bicycle and ... riding it across the road.*
BEG...VSS#|Mpo|sit§-on-bicycle...EO&|on-bicycle§-Mmi (Mmv+Tbo in CI line)
- *She is with a bicycle and ... going across the road.*
BEG...EO& |with-bicycle(inst)§-Mmi (Tbo in CI line)
- *The bag is attached to him via a rope, ... he’s pulling it up.*
BEG...VSS#|be§-attached-Mmx...EO&|via§-rope(inst)§-Mmx (Cse+Mac+Tve in CI line)
- *Elle a des patins à roulettes et ... elle traverse la route.*
BEG...VSS#|avoir§-poss-des-patins-à-roulettes...EN1/1&|patins-à-roulettes§-Mmi (Tbo in CI line)
- *Il y a une dame en vélo, ... elle avance [...].*
BEG... VSS#|y-avoir§-pres-loc-en-vélo...EO&|en-vélo(inst)§-Mmi (Tpa in CI line)
- *Il y a un nageur sur l’herbe ... et il arrive à une rivière.*
BEG... VSS#|y-avoir§-pres-loc... EN&|nageur§-Mmi (Tbo in CI line)
- *Da ist ein Jogger ... und er rennt über die Straße.*
BEG...VSS#|sein§-pres...EN&|jogger§-Mmi (Mmv+Tbo in CI line)

5) Screen references

When the computer screen (used to present the stimuli) is explicitly mentioned in a prepositional expression, the indication *écran* (French for computer screen) appears after the §- sign.

- *Der Mann läuft in den Bildschirm.* = EA1/1&+Tbo|in+NP§-écran
- *He is running across the screen* = EP1/1&+Tbo|across+NP§-écran
- *Il sort de l’écran* = EP1/1&+Tbo|de+NP§-écran

6) Non-spatial values

Certain non-spatial values of satellites and expressions other than the verb are indicated in parenthesis after the form used. The list below illustrates the most frequent cases.

Agent (agt): when a satellite-like expression expresses the agent in a passive utterance

- *be pushed by Hoppy* = EO...|by(agt).

Instrument (inst): For cases where an expression indicates an instrument used to carry out and/or provoke a displacement. In these cases, the instruments used are indicated by the annotation (inst) at the end. At the same time, the codes *Mac* or *Mmv* before the bar indicate the Manner of the displacement (provoked or spontaneous, respectively). One exception concerns parts of the body (hands, feet, claws, etc.), which are not annotated by the indication (inst) even when they contribute to the expression of Manner.

- on bicycle, on skates, on shoes = EO...&+Mmv|on-shoes(inst)
- *il tire le sac avec une corde:* = EO...&+Mac|avec-corde(inst)
- *elle avance en vélo / sur un vélo:* = EO...&+Mmv|en-vélo(inst)
= EO...&+Mmv|sur-vélo (inst)
- *il grimpe avec ses pattes:* = EO...&+Mmv|avec-pattes

- *il glisse sur ses pieds:* = EP...&+Loc|sur-pieds
- *sie fährt auf/mit einem Fahrrad über die Straße.* = EO...&+Mmv|auf-fahrrad(inst)

Comitative (com): Used to indicate the addition of an element relevant for the stimulus and the presence of which is expressed without any other semantic content:

- *He is on the hill with a bag:* = EO...&|avec-sac(com)

7) Multiple values

It is useful to note that a given expression can have different values depending on the context. Some examples:

entlang ('along'):

- ED&+Loc|auf+NP...EX&+Tpa|entlang *er geht auf der Strasse entlang* (Adverb)
- EA&+Tpa|NP+entlang *er geht die Strasse entlang* (Postposition)

By, with the help of, thanks to (and its translations):

Code EP if it is a spatial preposition:

- *passer **par** la porte* = EP...&+Tpa|par+NP

Code EO when *by* introduces the agent in a passive phrase, without indicating the semantic content:

- *be pushed **by** Hoppy* = EO...|by-Hoppy(agt)

Code EO if instrumental, by indicating 'Mac' as semantic content, and (inst) following the form of the expression:

- *tirer **par** une corde* = EO...&+Mac|par-corde(inst).
- *mit **hilfe** einer Leine* = EO...&+Mac|with-help-of-rope(inst).

Other cases, not coded for content:

- *commencer **par** la droite:* =EO...|par(commencer).

With, avec, mit:

Code for an instrument

= EO...&+Mmv|avec-vélo(inst)
or EO...&+Mac|avec- vélo(inst)

- *hit with a stick, tirer avec une corde, die fährt mit dem Fahrrad über die Straße.*

Code EO when *with* has an accompaniment value 'com' or comitative:

- *with car, with chair, with hobby horse* = EO...|with-chair(com)
- *Popi arrive avec une balle/ avec un sac sur son dos*
= EO...|avec-balle(com).
- *Popi überquert mit dem Schaukelpferd die Straße* ('Hoppy crosses the road with his wooden horse').
- *Ein Rad und Popi überquert damit die Straße* ('A wheel and Hoppy crosses the road with it').

Expand this coding (com) to similar values

- *it's a house with a sloping roof.*

Code EO if the information is relevant for Path or Manner:

- *Il marche **avec lenteur*** = EO&+Mmv/avec-lenteur
- *Er bewegt sich **mit schnellen Schritten/typischen Affenbewegungen***
= EO&+Mmv/mit-schnellen-schritten

Specific cases: *Der ist mit.* ‘He is with.’ = EP mit(com) without semantic content.

wieder, dann, weiter, encore, again... :

In the majority of cases, elements such as *wieder/dann/weiter/encore/again* (EO) have an aspectual value that is not coded on lines concerning space (but see Section IIC for the coding of their temporal values). In German, the boundaries between spatial and aspectual values are sometimes fuzzy, however, and a conservative code that codes *wieder* as being semantically void is chosen: EO1/1&|wieder. There are some exceptions with regard to certain cases such as *weiter* ‘still/even further’ which more clearly have a double temporal and spatial value and for which the spatial value is coded:

- *Sie geht weiter* = EL&+Tpa|weiter.

8) Other points relating to coding of satellites specific to a given language

• **French**

In French, *là* is always coded as EX&+Loc

Expressions such as *à quatre pattes* are coded as follows, according to the context in which they are used:

- *il est à quatre pattes:* VSS#+Mpo|être§-à-quatre-pattes
- *il se met à quatre pattes:* VMM#+Mpo|se-mettre§-à-quatre-pattes
- *il marche à quatre pattes:* VMM#+Mmv|marcher EO&+Mmv|à-quatre-pattes

In special cases of idiosyncratic use, particularly in children, direct and indirect object complements (COD) and (COI) may be coded:

- *il pousse avec le ballon* (= Il pousse le ballon) = EO...&|avec-ballon(COD)§-idioeo
- *il a roulé avec la bouée* (= Il a roulé la bouée) = EO...&|avec-bouée (COD)§-idioeo
- *Popi fait traverser la route à la poussette* = EO...&|à-poussette(COI)§-idioeo
- *Il fait grimper le toit à la voiture* = EO...&|à-voiture(COI)§-idioeo

• **German**

In principle, prepositional expressions that allow an alternation between dative and accusative, indicate a change of location when in the accusative (Tbo or Tgo). Nevertheless, if the entity serving as Ground corresponds to a large spatial area (for example, a meadow), the expression typically implies that the motion event has been carried out within this area. In these cases, the annotation §-area is added after the §- sign.

Er läuft über das Gras. ‘He walks on the grass.’ EA&+Loc|über+NP§-area

da§-initial

The adverb *da* ‘there’ in German is extremely frequent, particularly in children’s productions. In most cases, its semantic content is fuzzy, which makes it difficult to code a spatial value. It is almost always the case that *da* is used in the initial position of the utterance. In these cases, *da* is coded as a non-spatial expression (EO rather than EX) without semantic content, and its position is indicated after the §- sign with the annotation §-initial.

Da kommt ein Mann. = EO1/1&|da§-initial

Other cases:

Er ist da unten ‘He is down there.’ *da:* EX&+Loc, *unten:* EX&+Loc vs.
Der ist da hergerollt. ‘It rolled as far as there.’ *da:* EX&+Loc and *her:* EL#+Tdx

- **English**

Home

The term ‘home’ can be used as an adverbial expression (*go/be home*) or as a noun phrase with determiner (*go to his home*). This term is coded EX whenever it occurs. If it is used jointly with another satellite (*he went back home*), this is indicated after the §- sign that accompanies the verb.

Example	Verb	Satellite 1	Satellite 2	Ground
He is home	Be	EX&+Loc home	--	Grd=home
He is at home	Be	EX&+Loc at-home	--	Grd=home
He went home	Go	EX&+Loc home	--	Grd=home
He went back home	go§-back	EL&+Tdx back	EX&+Loc home	Grd=home
He went to his home	Go	EP&+Tgo to+NP	EX&+Loc home	Grd=home

All the way

The expression ‘all the way’ is mostly used with another spatial expression (*she went all the way up, all the way to the top, all the way across the road, all the way across to the other side, etc.*). It can also sometimes be used alone (*he ran all the way*). The coding for these varied cases is indicated in the table below.

Example	Satellite 1	Satellite 2	Satellite 3
<i>he ran all the way</i>	EX&+Tpa all-the-way	--	--
<i>he went all the way up</i>	EX&+Tpa all-the-way	EL&+Tve up§-all-the-way	--
<i>he went all the way up the tree</i>	EX&+Tpa all-the-way	EP&+Tve up+NP§-all-the-way	--
<i>he went all the way up to the top of the tree</i>	EX&+Tpa all-the-way	EL&+Tve up§-all-the-way-up-to	EP&+Tgo+Loc to-top-of+NP§-up-to
<i>he went all the way up to the tree</i>	EX&+Tpa all-the-way	EL&+Tve up§-all-the-way-up-to	EP&+Tgo to+NP§-up-to
<i>he went all the way to the other side</i>	EX&+Tpa all-the-way	EP&+Tgo+Loc to-other-side-of+NP§-all-the-way	--

Round/around

The terms *round* and *around* can have several distinct senses and be combined with other expressions. They can be coded as prepositional expressions (EP) or as particles (EL) according to whether or not they are accompanied by a noun, respectively.

Sense	Verb	Satellite
Manner of motion event = ‘roll’	go§-round/around	EL&+Mmp around
Indeterminate path = ‘all over’	go§-round/around	EL/EP&+Tpa around
Change of direction = ‘turn’	go§-round/around	EL&+Tpa around
Surround = dynamic surround	go§-round/around	EL/EP&+Tpa around
Surround = static surround	be§-round/around	EP&+Loc around
Be around = ‘here’	be§-around	EL&Loc around

There

Use the EX code to code *there*. When *there* is used with another element (for example, a particle or a preposition), code the two as separate elements by indicating the use of *there* after the §- sign on the other element, for example:

Static		Satellite 1	Satellite 2
<i>There</i>	He is there	EX&+Loc there	--
<i>X + there</i>	He is up/down/in/on/over there	EX&+Loc up-there EX&+Loc in-there EX&+Loc over-there	-- -- --
Dynamic			
<i>There</i>	He went there	EX&+Loc there	--
<i>X + there</i>	He went up/down there He went across there He went over, through, along there He went back there He went in/out there	EL&+Tve up§-there EL&+Tbo across§-there EL&+Tpa along§-there EL&+Tdx back§-there EL&+Tbo in§-there	EX&+Loc there EX&+Loc there EX&+Loc there EX&+Loc there EX&+Loc there
Prep + <i>there</i>	He went into, onto, from, into, out of... there	EP&+Tbo into§-there	EX&+Loc there

Examples:

He is down there VSS#|be§-loc..... EX&+Loc|down-there

He pushed it down there VMP#+Cse+Mac|push§-down...EL&+Tve|down§-there...EX&+Loc|there

2.1.9 Field 13: Periphery

This field specifies the type of phrase in which verbs and satellites have a role. Among other things, it is aimed at identifying certain obvious inter-propositional relationships, such as subordinate phrases providing information regarding the Path and/or the Manner in the periphery of the main verbal nucleus.

Independent or main clauses ('matrices') as well as subordinate clauses are identified. Coordinated clauses (*and*, *or*, etc.) are considered as independent.

Main or independent matrix clause:

PM0 A clause corresponding to a main or independent (non-subordinate) matrix clause that does not comprise subordinate elements in its periphery.

PMS A clause corresponding to a main clause (non-subordinate) to which one or several subordinate clauses are attached.

Subordinate clause:

PS| The type of subordination is indicated after the vertical bar (0+ppr, inf, *en+ger*, *pour+inf*, *sans+inf*, *rel*, *sondern*, *bis*, *weil*, *als*, *dass*, etc.). Subordinates with modal verbs and temporal-aspectual verbs are ignored (*he wants to cross*, *he continues to climb* are not segmented).

Examples:

- *There's a man running* = PS|0+ppr
- *he sees the ball fall down* = PS|inf
- *sans s'arrêter* = PS|sans+inf

- *de manière à descendre* = PS|de-manière-à+inf
- *als er die Straße überquert hat* = PS|als
- *da sieht man einen Strommasten [c], auf dem sich ein Nest befindet* = PS|rel

PMX| Special subordinate cases

This code is used for coding subordinate clauses that can be treated as main clauses for some analyses. In these cases, main and subordinate clauses are coded from a formal point of view as in the previous cases, except that the PS code is replaced by the PMX code, followed by the indication of type of subordination. The most frequent examples correspond to two types of case:

- 1) When the target is a subordinate of a main clause that is coded in another component or as irrelevant. Examples:

'He is in the water because he's swimming across the river' is coded as 'He swims across the river':

- *He is in the water* (= BEG)
- *because he's swimming across the river* (=PMX)

Il est malin parce qu'il est monté en grim pant is coded as 'Il est monté en grim pant':

- *Il est malin* (= IRR)
- *parce qu'il est monté* (=PMX)
- *en grim pant* (+PS|en+ger)

- 2) In complex presentative clauses (*C'est X qui...*, *There's an X who...*, *j'ai vu X qui*, *ich sah X inf...*). Examples:

C'est Popi qui fait descendre la valise en la poussant is coded as 'Popi fait descendre la valise en la poussant':

- *c'est Popi* = PMS (=BEG)
- *qui fait descendre la valise* = PMX|rel (=CIB)
- *en la poussant* = PS|en+ger

ich sehe dass der den Ball den Hügel runterrollt

- *ich sehe* 'I see', = PMS (=BEG)
- *dass der den Ball den Hügel runterrollt.* = PMX|dass (=CIB)

Examples for different types of clauses in each of the languages

Examples in English:

- *The caterpillar climbed up the leaf.* (PM0)
- *The cat licked up the egg yolk (PM0) / and walked away.* (PM0)
- *Hoppy walked up the roof (PMS) / pushing the present in front of him.* (PS|0+ppr)
- *There's a little mouse (PMS) / tiptoeing up to the table (PMX|0+ppr)*

Examples in French:

PM0

- *Il veut traverser* (PM0).
- *Il regarde (IRR), décide de monter (PM0) puis il est content (IRR).*

PMS + PS

- *Il traverse (=PMS) en nageant (=PS|en+ger).*
- *Il traverse (=PMS) parce qu'il va à l'autre bout (=PS|parce-que) en courant (=PS|en+ger) sans*

s'arrêter (=PS|sans+inf).

- *Et puis, il repart (PMS) d'où il est venu (PS|rel).*
- *Arrivé de l'autre côté de la rivière (PS|PP), il se remet à courir (PMS).*

Phrases with PMX

- *C'est Popi (PMS) qui fait descendre la valise (PMX|rel) en la poussant (PS|en+ger) = Popi fait descendre la valise en la poussant.*
- *C'est Popi (PMS) qui monte (PMX|rel) en poussant un ballon (PS|en+ger) qui roule jusqu'en bas*
- *(PS|rel). = Popi monte en poussant un ballon qui roule.*
- *Il y a un garçon (=PMS) qui traverse (=PMX) en courant (=PS)*
- *Il y a un garçon (=PMS) qui court (=PMX) et qui traverse (=PMX)*

Examples in German:

- *Ein Koffer und Popi wie sie den Hügel runtergehen. = Ein Koffer und Popi gehen den Hügel runter. (=PM0)*
- *Rutscht von links nach rechts durchs Bild und zwar auf einem zugefrorenen See. = Rutscht von links nach rechts durchs Bild [und zwar] auf einem zugefrorenen See (=PM0)*
- *Man sieht ein Eichhörnchen (=PMS), den Baum hochkrabbeln (=PMX|inf).*
- *Ein Junge überquert (=PMS) schwimmend (PS|PPR) den Fluss.*

2.1.10 Fields 14 and 15: Figures and Grounds

Figure and Ground coding

This field identifies the referents used in expressing location and motion events. This coding is carried out in connection with the verb (field 5), since only some verbs require the identification of these referents:

- **Figure** – the entity located and / or in motion.
- **Ground** – the entity or entities in relation to which the Figure is located, or moves.

FIELD 14: FIGURE

Figform(...)	Explicit Figure.
Figimp(...)	Implicit Figure.
Fi0 	Figure irrelevant.
Fix 	Coding difficulty.

FIELD 15: GROUND

Grd form(...)	Explicit Ground.
Grd imp(...)	Implicit Ground.
Gr0 	Ground irrelevant.
Gix 	Coding difficulty.

DETAILS

Figform(ref)	Explicit Figure and Ground.
Grd form(ref)	When an explicit form indicates the Figure or the Ground - noun or pronominal form (even if little informative: it, there, <i>il</i> , <i>y</i> , <i>là</i> , <i>da</i> , <i>etc.</i>).

a) **Form** – after the bar, the form used is indicated initially:

Figure: Indicate the noun used (*Hoppy, boy, balle, poussette, Affe, Tisch*, etc.) or PRO (for 'pronoun'), including when it concerns a zero pronoun (considered as an 'explicit' form, as opposed to implicit referents, see below).

Ground: Indicate the form used, even if it is a pronoun (*there, inside, road, y, dedans, drin*, etc.). Note: The notion of the zero pronoun does not apply to the Ground (see the imp code below) except in certain special cases, for example *il court après* (where the preposition implies the presence of a Ground with a zero pronoun).

b) **Referent** - the referent is then identified in parenthesis, where necessary and possible: If the form is nominal, do not put parenthesis after the form. For all less informative forms (pronouns, etc.), identify the referent from the stimulus, except where the reference is very vague. In this case, indicate (*indet*). In principle, cases coded (*indet*) are more frequent for Grounds than for Figures (see special cases below).

Fig|imp(ref) **Implicit Figure and Ground.**
Grd|imp(ref) The reference is entirely presupposed (no form).

Figure: For Figures, these cases are very rare (see special cases below) and do not include zero pronouns (considered as 'explicit' forms).

Ground: For Grounds, these cases concern those where no mention of a Ground is made at all. After the bar, indicate *imp* for the form, then infer the referent, or indicate (*indet*) if the reference is very vague.

Fi0|0 **Figure and Ground irrelevant.**
Gr0|0 Use the 'zero' sign (and not the capital letter 'O'). Put 'zero' after the bar.

Fix| **Coding difficulty (very rare in principle).**
Grx| Code used with certain verbs (see below and Appendices). Do not put anything after the bar.

Other codes to note after the form used and the referent in parenthesis:

-MP, -MU, -MN **Discursive status of referents in rare cases of less informative forms (i.e., pronominal or zero forms).**

For discursive analyses that are more qualitative, the informational status of the referent serving as the Ground can be indicated as follows: first (première) mention (MP), subsequent or ulterior mention (MU), no mention if the referent is never mentioned (MN). The discursive status of the referent is indicated in parenthesis (with a hyphen after the referent).

-MP *First Mention*
Less informative form for the very first mention of the referent;
-MU *Ulterior Mention*
Less informative form for a referent already mentioned in the subject discourse;
-MN *No Mention or 'never mentioned'*

Total absence of mention (in the utterance and the previous discourse of the subject).

Note: In principle, these codes only apply to the Ground, except in certain special cases (see below).

Special cases:

Special cases relate to utterances in which the referent's discursive status must be specified for Figures (in productions by young children, aphasics, L2 beginners, etc.):

<i>Y a une colline et il le pousse en bas.</i>	Fig PRO(ballon-MP):Grd imp(colline-MU)
<i>Y a une colline, il pousse et il s'en va.</i>	Fig imp(ballon-MN):Grd imp(colline-MU)
<i>Popi arrive, il pousse et il s'en va.</i>	Fig imp(ballon-MN):Grd imp(colline-MN)
Push down.	Fig imp(ballon-MN):Grd imp(colline-MN)
Climb up.	Fig imp(bear-MN):Grd imp(tree-MN)
Up tree.	Fig imp(bear-MN):Grd tree

Criteria for determining the presence/absence or the relevance of Figure and Ground

1. Real Figures and Grounds: Fig and Grd

Some utterances necessarily imply the reference to a Ground (coded Fi... and Gr...), even if it is not mentioned. The presence or absence of an explicit mention is not what constitutes the criterion for the relevance of these notions. Their relevance depends on semantic criteria: is this a question of a situation implying the location and/or the displacement (explicit or implicit)?

There will always be a Figure and a Ground (implicit or explicit) in the following cases:

(a) With a motion verb (voluntary or caused motion), with or without an explicit spatial expression:

VMM: *swim, run, walk, aller, se coucher, tomber, laufen, kommen, etc.*

VMP: *push, pull, take, mettre, enlever, ramener, schieben, ziehen, etc.*

(b) When there is an explicit spatial expression in the utterance, whatever the verb:

- with static verbs **VSS:** *be in/on, be there/here, etc.*
- with dynamic verbs **VXX:** *eat at the bottom of the tree, etc.*

2. Figures and Grounds Irrelevant: Fi0 and Gr0

In some utterances, the notions of Figure and Ground are irrelevant, for example, in existential or presentative utterances (*there's a cat, c'est le chat de la voisine*) in the absence of an explicit location. This is also the case for utterances containing VSS or VXX without location and without implicit displacement (e.g. eat, sleep, be-ered, *écraser, casser, déchirer, zeigen, klopfen, fressen*). Also include the non-spatial use of some verbs, coded 'vns' (*come(vns), walk(vns), go(vns), etc.*).

- *The girl has put on skates.* Fi0 Gr0
- *Les arbres sont verts.* Fi0 Gr0
- *Das Eichhörnchen scheint pink zu sein.* Fi0 Gr0

3. Questionable Figures and Grounds: Fix and Grx

Some utterances present Figures and Grounds with questionable relevance (coded 'Fix' and 'Grx'), for example, for some verbs such as 'open', 'close', 'carry', *rouvrir, trouver, suchen, lassen*. Use these codes, except where the utterance contains an explicit location. The questionable nature of the Figure and the Ground does not come from the fact that the referent is implicit or difficult to identify, but rather from the nature of the verb, the structure and/or the function of the utterance. With an explicit location, always code the Figure and the Ground (*He finds the honey in the hive*). In the absence of an explicit location, code 'indet' (*Il ferme la boîte*).

- *The squirrel goes to look for nuts [Fig] in the tree [Grd].* Fig Grd
- *L'écureuil va chercher des noisettes [Fix] and [Grx indet]* Fix Grx
- *Er macht die Garage auf [Fix] and [Grx indet]* Fix Grx

Additional distinctions

1. Figure Identification

Only code as Figure referents for which the motion or location is necessarily expressed by the verb. Do not code as Figure referents for which the motion is inferred by the context. For example:

With caused motion, the subject of the phrase is not the Figure:

- *Hoppy pushes the cartwheel.* Figure = cartwheel

With caused motion, referents that can be carried by the agent are not Figures:

- *Ein Korb und Popi überquert damit die Straße.* = Fig|Pro(popi)
- With static locations, referents that can be located together are not all Figures:
- *The bear is under the tree with the honey.* = Fig|bear

2. Important rule for Figure/Ground pairs

The codes in fields 14 and 15 constitute ‘symmetrical pairs’:

- if Figure = Fig| => Ground = Grd|, and vice-versa.
- if Figure = Fi0| => Ground = Gr0|, and vice-versa.
- if Figure = Fix| => Ground = Grx|, and vice-versa.

3. Several Figures or Grounds

When several Figures or Grounds are mentioned in the same clause (corresponding to several satellites, for example), indicate the different referents after the bar with the + sign between them, arranging them in the order in which they are mentioned. Examples:

- *He rolls the ball into the cave in the forest* Grd|cave+forest.
- *Elle marche sur le trottoir jusqu’à la maison.* Grd|trottoir+maiso
- *Popi schleppt den Stuhl hinter sich in die Höhle.* Grd|PRO(Popi)+hö
- *Ein Koffer und Popi gehen den Hügel runter.* Fig|koffer+Popi

4. Form coding

(a) **Relative pronouns:** Indicate ‘PRO’ as for a third person pronoun.

The ‘rel’ code appears in the periphery coding (see field 13).

(b) **Dislocations:** Indicate the name only: *la femme elle fait la bicyclette* = Fig|femme

(c) **Other forms:**

Certain forms coded as EX in satellites can appear in Ground coding. For example (also see examples in Table 6 below):

- In English, the Ground can be pronominalised by forms like *inside*.
- In French, the Figure can be pronominalised by a personal pronoun or by *en*, the Ground by *y* or by *où*.
- In German, the Ground can be pronominalised by *da* or forms composed with *da*.

French – *il y va en grim pant*

English – *He sits inside*

German – *Er ist da*

Tables 6 and 7 provide an overview of the various possible codings of Figures and Grounds as discussed above, including possible combinations of Figures and Grounds and types of satellites and Grounds.

Table 6: Examples of Figure and Ground coding in English, French and German

French		
<i>Popi descend.</i>	<i>Fig/popi</i>	<i>Grd/imp(colline-MN)</i>
<i>Il traverse la rue.</i>	<i>Fig/PRO(popi)</i>	<i>Grd/rue</i>
<i>(Il a un pneu) et il va dedans avec le pneu</i>	<i>Fig/PRO(popi)</i>	<i>Grd/dedans(grotte-MP)</i>
<i>Il voit une grotte et il y va.</i>	<i>Fig/PRO(popi)</i>	<i>Grd/PRO(grotte-MU)</i>
<i>Il est chez lui.</i>	<i>Fig/PRO(popi)</i>	<i>Grd/PRO(Popi)</i>
<i>(Il a une poussette) et il va là et là.</i>	<i>Fig/PRO(popi)</i>	<i>Grd/là(indet)</i>
<i>Là, il a une chaise.</i>	<i>Fig/PRO(popi)</i>	<i>Grd/là(indet)</i>
<i>(Il traverse) et il arrive de l'autre côté de la rue.</i>	<i>Fig/PRO(jogger)</i>	<i>Grd/rue</i>
<i>Il traverse et (il arrive de l'autre côté.)</i>	<i>Fig/PRO(jogger)</i>	<i>Grd/imp(rue-MN)</i>
<i>Il est content.</i>	<i>Fi0/0</i>	<i>Gr0/0</i>
English		
<i>(The girl is in the forest) and skates across the lake.</i>	<i>Fig/PRO(girl)</i>	<i>Grd/lake</i>
<i>(The girl comes to the lake) and skates across.</i>	<i>Fig/PRO(girl)</i>	<i>Grd/imp(lake-MU)</i>
<i>(The cat sees the egg) and climbs up to the nest.</i>	<i>Fig/PRO(cat)</i>	<i>Grd/nest</i>
<i>(The squirrel runs up the tree) and back down.</i>	<i>Fig/PRO(squirrel)</i>	<i>Grd/imp(tree-MU)</i>
<i>The bear climbs up (and comes back down).</i>	<i>Fig/ bear</i>	<i>Grd/imp(tree-MN)</i>
<i>(The bear climbs up) and comes back down.</i>	<i>Fig/PRO(bear)</i>	<i>Grd/imp(tree-MN)</i>
<i>(Hoppy has a balloon) and he pushes it down the hill.</i>	<i>Fig/PRO(ballon)</i>	<i>Grd/colline</i>
<i>(Hoppy has a balloon, he pushes it and) it goes there.</i>	<i>Fig/PRO(ballon)</i>	<i>Grd/imp(colline-MN)</i>
German		
<i>(Der Junge kommt und) schwimmt durch den Fluss.</i>	<i>Fig/PRO(junge)</i>	<i>Grd/fluss</i>
<i>Er hält die Schubkarre hinten.</i>	<i>Fig/PRO(popi)</i>	<i>Grd/imp(popi-MU)</i>
<i>Der Ball kickt die Würfel um.</i>	<i>Fig/würfel</i>	<i>Grd/imp(indet)</i>
<i>(Der Mann schwimmt und) verschwindet wieder.</i>	<i>Fig/PRO(mann)</i>	<i>Grd/imp(indet)</i>
<i>Die Flasche geht kaputt.</i>	<i>Fi0/</i>	<i>Gr0/</i>

Table 7. Examples of Satellite and Ground coding

Example	Satellite coding	Ground coding
<i>He walks down from the sanddune</i>	EP+Tso from+NP	Grd sanddune
<i>He walks up</i>	EL+Tve up-z	Grd imp(roof)
<i>He walks inside</i>	EX+Loc inside	Grd inside
<i>There is a caterpillar there</i>	EX+Loc there	Grd there
<i>Il va dedans.</i>	EX+Loc dedans	Grd dedans
<i>Il est là.</i>	EX+Loc là	Grd là
<i>Il est au village.</i>	EP+Loc à+NP	Grd village
<i>Elle est arrive là</i>	EX+Loc là	Grd là
<i>Il va jusque là (-bas)</i>	EP+Tgo+Loc jusque-là	Grd là(-bas)
<i>Elle est partie de là</i>	EP+Tso+Loc de-là	Grd là
<i>Elle est passé par là</i>	EP+Tpa+Loc par-là	Grd là
<i>Il est allé vers là, vers là-bas</i>	EP+Tdi+Loc vers-là	Grd là
<i>La maison se situe dans la forêt.</i>	EP+Loc dans+NP	Grd forêt
<i>...jusqu'ici, d'ici, par ici...</i>	EP+Tgo+Loc jusque-ici	Grd ici
<i>Il est sur la colline</i>	EP+Loc sur+NP	Grd colline
<i>Il est dessus</i>	EX+Loc dessus	Grd dessus
<i>Il met la roue dans la grotte</i>	EP+Loc dans+NP	Grd grotte
<i>Il met la roue dedans</i>	EX+Loc dedans	Grd dedans
<i>Il va à l'arbre</i>	EP+Loc à+NP	Grd arbre
<i>Il est à droite</i>	EP+Loc à-droite	Grd droite
<i>Il est à la droite</i>	EP+Loc à-la-droite	Grd la-droite
<i>Il est à ma droite</i>	EP+Loc à-ma-droite	Grd ma-droite
<i>Le chat monte</i>	EZ ...	Grd imp...
<i>Il est tombé par terre</i>	EX+Loc par-terre	Grd par-terre
(an object falls down) <i>Ça tombe.</i>	EZ ...	Grd imp...
(an object falls on a table) <i>Il tombe.</i>	EZ ...	Grd imp...
<i>Elle arrive au bas de la colline</i>	EP+Loc au-bas-de+NP	Grd le-bas-de-colline
<i>Elle arrive en bas</i>	EP+Loc en-bas-z	Grd en-bas(colline)
<i>Elle arrive au sommet de la colline</i>	EP+Loc à+NP	Grd sommet-de-colline
<i>Il est en haut de l'arbre</i>	EP+Loc en-haut-de+NP	Grd en-haut-de-arbre
<i>(Il y a un arbre.) Il monte en haut</i>	EP+Loc en-haut-z	Grd en-haut(arbre)
<i>Il est en haut</i>	EP+Loc en-haut-z	Grd en-haut(-arbre)
<i>Er geht da rein.</i>	EX+Loc EL+Tbo+Tdx rein	Grd da
<i>Er ist drüber gegangen</i>	EL+Tbo+Tdx+Loc drüber	Grd da
<i>Er läuft rüber.</i>	EL+Tbo+Tdx rüber	Grd imp(straße)
<i>Er läuft dort.</i>	EX+Loc dort	Grd dort

2.1.11 Field 16: Comments

Various comments are added in brackets at the end of the line, if necessary. If there are no comments, click on {} in the coder menu all the same to produce empty brackets. Some of these annotations can be invented, as necessary. However, most are recurrent and have been standardised (see section A below). In some cases, specific syntax is used in the brackets in field 16 (see section B below). The annotations in this field are different (but sometimes redundant) from annotations appearing inside other fields (after the §- signs). The list of the latter is summarised in section C below.

Common, standardised comments in Field 16

{autorep}	Immediate and completely identical repetition by the subject with or without experimenter intervention:
*CHI:	<i>he has left.</i>
*EXP:	<i>Yes, he has gone back to the house.</i>
*CHI:	<i>he has left.</i>
{repetition}	Immediate and identical repetition by the subject from an utterance by another speaker and in the absence of an intervening utterance:
*CHI:	<i>he has left.</i>
*EXP:	<i>Yes, he has gone back to the house.</i>
*CHI:	<i>he has gone back to the house.</i>
{embed}	Embedded clause: <i>The cat that had climbed on the post [c] made the egg fall. [c]</i> <i>the cat made the egg fall</i> = Main clause coding = IRR <i>that had climbed on the post</i> = Relative coding {embed} = CIBLE = PMX
{onomat}	Onomatopoeia (e.g.: <i>fffft</i> , <i>vroum</i>).
{incomplet}	Incomplete utterance (includes interruptions): <i>Il monte sur...</i>
{inaud}	Inaudible utterance.
{formule}	Example: Please; <i>Tiens!</i> Also use with (vns) verbs: <i>Got it! J'arrive pas.</i>

The following types of construction are considered as being idiomatic in German: *vor/hinter sich her*. In these cases, the prepositional expression *vor/hinter+sich* accompanied by the particle *her* expresses a Path by specifying the location of the motion. The code *Tpa+Loc* is used for the semantic content, and the annotation *§-formule* indicates that it is an idiomatic expression.

- *Er zieht den Koffer hinter sich her.* = EP1/1&+Tpa+Loc|hinter-sich-her§- formule
- *Er schiebt den Koffer vor sich her.* = EP1/1&+Tpa+Loc|vor-sich-her§-formule

{question} The subject's utterance is a question.

{autocor-x/y...} Auto-correction in the utterance: one element is changed into one or several others. In principle, the last element is the one that will be taken into account in the analysis. If possible, specify the corrected element or the minimum amount, to enable more complex reformulations to be found:

- {autocor-on/in} indicates that the subject said *on* first, then *in*
- {autocor-courir-en-traversant/traverser-en-courant} indicates the reformulation indicated below:
- *Il court en traversant la euh non il traverse la route en courant.*
- *Er rutscht oder gleitet über das Eis.* {autocor- rutschen/gleiten}

{...x2} Element mentioned more than once in the utterance:
'Here it's white here' = {herex2}

{idio...} If an element in the participant response is **idiosyncratic** coders can use the comment 'idio' followed by the element that is idiosyncratic. For example, 'idioutt' could indicate that the whole utterance is idiosyncratic, i.e., the general structure could be idiosyncratic in use. Or 'idiowo' could indicate that the word

order is idiosyncratic. Table 8 below shows some further examples including the starter code ‘idio’. Such comments are likely to be more frequent in child data, L2 learner data and data from aphasics.

Example:

moving caterpillar up on [=? stem]

Interpreted as ‘the caterpillar is moving up on...’

...VMM#|move\$-idionoaux:AMB:...Fig|caterpillar:Grd|stem:{idiowo}

{ambigvm}	Ambiguity between a future tense marking or motion verb <i>Je vais manger tout à l’heure...</i> (temporal, unambiguous) <i>Je vais à la cuisine...</i> (spatial, unambiguous) <i>Je vais manger à la cuisine</i> (ambiguous)
{GBE}	Indication that a lexical item has been given by the experimenter (Given By Experimenter)

Syntax in the bracketed sections of Field 16

Auto-corrections {autocor-dans/sur} where *sur* corresponds to the subject’s final choice.

+ If several comments are necessary, they are added with the ‘+’ sign
{autocor-il/la+idioutt}

Annotations that can appear in fields other than Field 16

Some general annotations in field 16 are partially redundant with other annotations placed after the relevant elements in the utterance (verb, satellite, etc.). We provide an alphabetical list of such codes in Section 2E with suggestions for the coding of L2 learner data, as many of these instances are typical for such participants.

2.2 Synthesized coding lines (%sum)

After the %cod lines for a given item, a %sum line gives a synthetic summary of the global architecture of the responses, their degree of complexity, and the distribution of semantic information in and between the utterances. This line includes the utterance considered as ‘principle target’, as well as all the other productions around this target. It can also run in parallel with coding gestures. The majority of the coding options are entered automatically with a coding menu, but others must be entered by hand.

2.2.1 General format of the line

This line always begins with the %sum: sign (‘summary’), followed by a tab.

There are then several fields with sub-fields separated by delimiters.

- **Delimiters:** The colon delimiter (:) separates the fields, other delimiters (&, #, \$) separate the sub-fields.
- **Fields 1 and 2:** Type of response and item name.
- **Field 3:** Global structure.
- **Fields 4-11:** Four components, each comprising the component name, number and its semantic content

Syntax:

%sum:[tab]\$Item-type:item-name:global_{tot}:comp_a:info&info#word\$-word:comp_b:info&info#word\$-word:comp_c:info&info#word\$-word:comp_d:info&info#word\$-word
 where $tot = a + b + c + d$

2.2.2 Fields 1 and 2: Identification of Item and Component types

In field 1, the \$ sign at the beginning of the line is followed by a code for the type of motion and/or the component of the scenario coded. Please note that IRR lines in the set of %cod lines do not reappear in the set of %sum lines.

\$M = ascending motion in the voluntary motion events series (Voluntary motion cartoons vertical).

\$D = descending motion in the voluntary motion event series (Voluntary motion cartoons vertical).

\$0 = motion component for all the other items (Hoppy; horizontal across-type motion in voluntary motion cartoons).

\$B = ‘beginning’ component (all cartoons).

\$R = ‘end’ component (all cartoons).

Field 1 is followed by the abbreviation of the item name in field 2 (for example XSOU). In the same way as is done for %cod coding lines, zero lines are inserted if the subject has not described one of the components (end, beginning, etc.). For the complete list of item names, see Appendix xxx in Volume 2.

*SUI: The little mouse climbs up the table (M), gets the cheese and slides back down (D) and tiptoes off (R).

%sum: \$B:XSOU:GXX0:BZ:zi#zm:BZ:zi#zm:BZ:zi#zm:BZ:zi#zm
 \$M:XSOU:....
 \$D:XSOU:....
 \$R:XSOU:....

*SUI: La dame traverse la voie ferrée en vélo (0) et elle continue sur l’autre chemin (R).

%sum: \$B: XVEL:GXX0:BZ:zi#zm:BZ:zi#zm:BZ:zi#zm:BZ:zi#zm
 \$0:XVEL:....
 \$R:XVEL:....

*CHI: Popi steht an einer Straße (B) und er zieht das Schaukelpferd rüber (0).

%sum: \$B:tgt_cherue:....
 \$0:tgt_cherue:....

\$R:tgt_cherue:GXX0:RZ:zi#zm:RZ:zi#zm:RZ:zi#zm:RZ:zi#zm

The general coding format is ultimately the same for all %sum lines, but the codes are more detailed for some lines (\$M, \$D, \$0) than for others (\$R, \$B). We start from the principle that four sub-components per %sum line must include all instances of possible Figures (from a minimum of one to a maximum of four). If no sub-components and/or fewer than four sub-components are expressed in a given response, line coding continues using the corresponding ‘empty’ codes (MZ:zi#zm, see below).

The coding order corresponds to the chronological order of the utterance. For example, the ‘target’ component can be produced first, among other components, or at the end of the response.

\$ Type:name-item:global_{tot}:**target**_a:info#word:other_b:info#word:other_c:info#word:other_d:info#word

\$ Type:name-item:global_{tot}:other_a:info#word:other_b:info#word:**target**_c:info#word:other_d:info#word

\$ Type:name-item:global_{tot}:other_a:info#word:other_b:info#word:other_c:info#word:**target**_a:info#word

2.2.3 Field 3: Global structure

Responses are regrouped into several principle types in accordance with the components on the other coding lines that have been coded in field 2. Among the codes proposed, responses can be distinguished according to two dimensions: ‘tight’ vs. ‘loose’ (compressed or distributed information), and ‘simple’ vs. ‘complex’. The indication _{tot} gives the total number of propositions contained in a given response, including those of the target component (C...) and those of other possible components that are potential targets (M...).

GTS_{tot} ‘global tight simple’: all the information expressed is brought together in one single clause that necessarily relates to the ‘target’ component.

GTC_{tot} ‘global tight complex’: all the information expressed is contained in a single utterance, which relates to the ‘target’ component, but comprises at least one subordinate clause.

GLS_{tot} ‘global loose simple’: the information is divided over several clauses, some of which are coordinated, comprising the ‘target’ component (codes beginning with the letter C) and one or more parallel components (codes beginning with the letter M).
Contrary to GLC cases, the target (C) component here only contains one clause (no subordinate clauses), whatever the nature of M components might be.

GLC_{tot} ‘global loose complex’: the same as GLS, except that the target (C) component contains more than one clause, whatever the nature of M components might be.

GZZ No target response for \$M, \$D and \$0 lines: this code is used if there is no part of the response that can really be considered as relevant to the target, even potentially. This code should be rare, and is mainly reserved for cases where the subjects (particularly young children) only produce static and/or off-subject responses (largely corresponding to %cod lines: CIB/NO).

GXX_n, GXX0: Temporary codes for \$B and \$R lines:

Researchers may decide to code the ‘global structure’ dimension only on \$M, \$D, and \$0 lines – that is, lines that record ‘main target’ (C) or ‘potential target’ (M) components. ‘Temporary’ placeholder codes have therefore been devised for the other lines (\$B and \$R):
GXX_n = if the response contains a beginning and/or an end;
GXX0 = if the response contains no beginning and/or end.

2.2.4 Fields 4 to 11: Content of Responses

\$M, \$D, \$0 lines: Splitting up targets into C and M components

Format detail: ...:comp_a:info&info#word§-word:...

For each component, its type and the number of clauses it contains (comp_a) are indicated, followed by the semantic information expressed (info), then the words used by the subject (word).

1) Components and number of clauses

Components always begin with either C (target) or M (potential target). There cannot be more than one code beginning with C on any given %sum line. In the event of a non-response (an off topic or missing response), the 'empty' codes CZ (initially) and MZ (three times following CZ) are used. For beginnings and ends (if coded), use the BZ (beginning) and RZ (end) code in cases where no utterance has been produced.

Non-response to the target: \$0:item:GZZ:CZ:zi#zm:MZ:zi#zm:MZ:zi#zm:MZ:zi#zm

Non-response for BEG: \$B:item:GZZ:BZ:zi#zm:BZ:zi#zm:BZ:zi#zm:BZ:zi#zm

Non-response for RSF: \$R:item:GXX0:RZ:zi#zm:RZ:zi#zm:RZ:zi#zm:RZ:zi#zm

The list below shows the possible codes for C and M components. Note that the letter following C is nearly always the same as that following M, except in the following cases:

CB, indicating that there is only one possible target, with which there must be no M apart from MD and MQ;

MD and MQ, which can be combined with any C;

MZ, which can be combined with any C, including with CZ (in the event no response is given).

These codes and the number of clauses that follow must follow certain rules of internal coherence with regard to the C/M relationship and to the global structure.

Whatever the global code might be:

There is always just one C component and three M components.

With the global GTS code $_{tot} = 1$ always (by definition).

GTS1 is always accompanied by CB1 and conversely, CB1 by GTS1 There is never an M.

He walks up the hill. / Il marche dans la rue / Elle monte jusqu'en haut...

.....:GTS1:CB1:...

With the global GTC code $_{tot} > 1$ always.

GTC is always accompanied by CB. There is never an M.

'He walks up the hill pulling the car up'...

.....:GTC2:CB2:...

Il monte en marchant et en tirant le sac...

.....:GTC3:CB3:...

With the global GLS code $_{tot} > 1$ always.

C can be of different types (CP, CS, etc.) and can contain just one clause. It is always accompanied by one or more M components.

He crosses the road and runs.

.....:GLS2:CP1:....:MP1:...

Il court ... il passe sur la route et il traverse jusqu'à l'autre bout.

.....:GLS3:MP1:....:MP1:....CP1...

With the global GLC code $_{tot} > 1$ always.

C can be of different types (CP, CS, etc.) and must contain more than one clause. It is always accompanied by one or more M components.

Il traverse la rue en tirant son cheval [...] et il marche...

.....:GLC3:CR2:....:MR1:....

Il grimpe en s'agrippant à l'arbre, il monte tout en haut...

.....:GLC3:CR2:....:MR1:....

1) Semantic content

After the colon, the semantic content (info) of the output is indicated before the '#' hash sign. Codes are abbreviations of the codes used on the %cod line.

C	Cause
T	Path (Trajectoire)
M	Manner
A	Action causing a motion event (this code relates to the Mac code in %cod lines)
L	Location (for satellites only)
Z	A form is used and has been coded, but does not contain any of the above information
zi	No form has been used (and therefore no information is expressed)

- *he rolls the rubber ring up the roof* = CTM

After the '#' hash sign, the relevant words used by the subject (*word*) are indicated.

The main verb (in infinitive or quotation form as appropriate for the language) is indicated on the one hand, and the other elements (satellites and periphery) on the other. A minimum amount of information concerning the Ground is also indicated when it has been mentioned by the subject. If the subject has not produced any satellites or information in the periphery, code zm.

NB. The 'zm' code in this part of the line (after the #) always goes with the 'zi' code for semantic content (before the #, see above).

As far as possible, the coding order corresponds to the chronological order of the utterance (but with exceptions). If multiple semantic types of information are expressed in the same linguistic means, use the standard order for combining this semantic information (cf. below).

- *Il traverse en courant* = T&M#traverser\$-en-courant
- *Il court pour traverser* = M&T#courir\$-pour-traverser

2) Delimiters

Except for the colon delimiter (:), which separates fields, the following delimiters are used for fields 4 to 11.

The hash (#) sign:

The hash sign separates semantic information (*info*) from words used by the subject (*word*).

The ampersand (&) sign:

This sign separates the semantic information contributed by different elements of the clause, in particular by the main verb, on the one hand, and by the other elements (satellites, subordinate clauses), on the other. There is always just one '&' in the semantic information sub-field.

Examples:

he walks (M) on all fours (M) = M&M

Il traverse (T) en courant (M) = T&M

Er klettert (T) auf dem Baum (L) = T&L

The ‘paragraph’ (§) sign:

This sign separates the words used by the subject corresponding to the verb versus other elements (satellites and periphery). There is always just one § sign in the sub-field of words followed by a hyphen.

Examples:

run§-*up-tree*
run§-*along-on-road*
marcher§-*à-quatre-pattes-dans-rue*
monter§-*sur-arbre-jusqu’au-trou*
descendre§-*de-l’arbre-en-glissant*
springen§-*am-baum*
gehen§-*wieder-herunter*
laufen§-*weg*

The hyphen (-) sign:

Use the hyphen for all links enabling the avoidance of a space on the coding line, particularly to link the words used by the subject after the hash (#) sign. In particular, use the hyphen when multiple types of semantic information are brought by different elements within the VERB or OTHER fields (except when this information is lexicalised, see below). In these cases, indicate all the types and sub-types of semantic information (e.g.: T-M or T-T or M-M or T-T-M).

Examples:

He climbs (M) up the hill (T) all the way to the top (T) = M&T-T
Elle monte (T) sur l’arbre (L) jusqu’ en haut (T) en utilisant ses pattes (M) = T&L-T-M
Il fait (C) monter (T) le ballon sur la colline (L) = CT&L

Do not use the hyphen if certain semantic information is lexicalised together in an identical lexeme, and always give the codes in a fixed order, indicated in the list below. Examples:

He pushes the ball: to push = CA
He rolls the ball: to roll = CM
He rolled the balloon down the hill = CM&T#roll§-down-hill
Il monte la valise: monter = CT
Il grimpe: grimper = TM
Il pousse le ballon jusqu’ en bas avec les mains = CM&T-M#pousser§-jusqu’ en-bas-avec- mains
Er schiebt den Tisch in die Grotte rein = CA&T-T#schieben§-in-grotte-rein

§B and §R lines: Complementary information arising from ‘beginning’ and ‘end’ components

In cases in which the ‘beginning’ and ‘end’ components (BEG and RSF on the main %cod code lines) provide relevant complementary semantic information, these are coded according to the same format, except that the coding rules are less detailed and restrictive.

Different relevant elements at the beginning or the end are listed on the corresponding lines (§B and §R, respectively), separated by the ‘colon’ (:) sign. Each element starts with the code ‘BGn’ or ‘RFn’, followed by relevant semantic information (separating the main verb and the rest with the & sign), followed by the ‘#’ sign, then followed by the the corresponding words produced by the participant (separating the main verb and the rest with the § sign). If no part of the response corresponds to the ‘beginning’ and ‘end’ components, the BZ and RZ signs are used, as well as the same ‘empty’ codes that are provided for the \$M, \$D and \$O (BZ:zi#zm and RZ :zi#zm) lines.

Coding non-responses:

%sum: [tab] §B:GXX0:BZ:zi#zm:BZ:zi#zm:BZ:zi#zm:BZ:zi#zm

%sum: [tab] \$R:GXX0:RZ:zi#zm:RZ:zi#zm:RZ:zi#zm:RZ:zi#zm

Example response:

*He gets out of his pram (BEG), he crawls on the pavement (BEG) and walks towards a crossing (BEG), **he goes on walking (MIP) and he crosses the road (CIP)**, and he crawls off again on the other pavement (RSF).*

The regular text in the example above shows the relevant information coming from the ‘beginning’ and ‘end’ components (main %cod coding lines BEG and RSF, %sum \$B and \$R lines).

The **boldface** text shows the main components (codes M and C on the main %cod lines, codes \$M, \$D and \$O on the %sum lines).

Table 8: Clause components and coding

Main %cod coding	%sum coding	
BEG (beginning)	\$B	The baby gets out of his pram
BEG (beginning)	\$B	He crawls (M) on the pavement (L)
BEG (beginning)	\$B	And walks (M) towards the crossing (T)
MIP (potential target)	\$O	He goes on walking (M)
CIP (target)	\$O	And he crosses the road (TM)
RSF (end)	\$R	And he crawls off again on the other pavement

%sum: \$B:XBEB:GTC3:BG1:Z&T#get\$-out-of:BG1:M&L#crawl\$-on:pavement:BG1:M&T#walk\$-towards-crossing:BZ:zi#zm
 \$O:XBEB:GLS2:MIP1:M&zi#walk\$-zm:CIP1:T&zi#cross\$-zm:MZ:zi#zm:MZ:zi#zm
 \$R:XBEB:GTS1:RF1:M&T-L:crawl\$-off-on-pavement:RZ:zi#zm:RZ:zi#zm:RZ:zi#zm

Problems linked to the coding of response types: Responses to questions posed by the experimenter

The synthetic lines take account of the response type, which is coded on %cod lines according to the general directives that distinguish the following cases (see %cod lines in Field 3):

- LR – output that is entirely spontaneous (absence of questions).
- LG – responses to so-called ‘general’ questions.
- LS – responses to so-called ‘specific’ questions.

The principle of this coding consists of retaining in the analysis spontaneous productions only, or productions in response to fairly general questions that have not ‘contaminated’ the participants’ production unduly. This coding is fairly strict in that it limits the LG code to the most general questions (And then? *And so?*) treating all the rest as LS. On the %sum lines, the responses to specific questions are coded MQS in the field corresponding to the component (see Field 2 of the %cod lines).

1) Simple case: Totally spontaneous response without a question: In this case, the coding continues as indicated above.

- *Il pousse la roue = LR / et il traverse = LR*
 GLS2 CR1=CA&zi#pousser-roue\$-zm / MR1= T&zi#traverser\$-zm

2) More complex cases, where the experimenter's questions intervene:

In the event of a question from the experimenter, the coding indicates whether the question is specific or general. The analysis that follows does not take account of responses to specific questions, while responses to general questions are considered as spontaneous responses. In synthetic %sum lines, LS entries become MQS entries.

- *He is pulling the pram = LR/and then? /pulling it across the road = LG*
GLS2 MR1= CA&zi#pull-pram§-zm/CR1=CA&T#pull§-across-road
- *Il traverse = LR/ Et après?/Il pousse la roue = LG*
GLS2 MR1= T&zi#traverser/CR1= CA&zi#pousser-roue§-zm
- *He pushes the wheel across the road = LR/And what does the wheel do? /It turns = LS*
GTS1 CB1=CA&T#push- §- wheel-across-road / MQS1=M&zi#turn§-zm
- *Il pousse la roue = LR/Et il fait quoi avc la route? /Il traverse = LS*
GTS1 CB1= CA&zi#pousser-roue§-zm / MQS1= T&zi#traverser§-zm
- *Er überquert die Straße mit ein Rad= LR/Und was macht er mit dem Rad? /Er schiebt es = LS*
GTS1 CB1= T&Z#überqueren§-mit-rad/MQS1= CA&zi#schieben-rad§-zm

Special cases:

When there is a more or less exact repetition of a given response (coded MID in the %cod line) after a general question (LG), MD is selected in the %sum line (ignoring MQG), and the total number of utterances is not added in the global structure. Examples:

```
%cod → He ran to the tree (LR, CIB), he uh... (EXP: And then?) He ran to the big tree (LG, MID)
%sum → ...GTS1...CB...MD...
%sum: GTS1:CB1:T&M #monter§-en-courant:MD1:M&zi #courir§-zm
%sum: GTS1:CB1:T&M #traverser§-avec-roue-en-courant:MD1:T&Z#traverser§-avec-roue
```

Additional information (example: negation) from the %cod line can be indicated in the %sum line with hyphens:
It did not move...

```
%cod:...RSF:...VMM#|move§-neg:.....
%sum:...RF1:Z&zi#move-neg§-zm:...
```

Examples from the Voluntary Motion series:

*SUIJ: *the baby [c] crawling across the path and to the right [c] and he crawled onto the path and onto side [c].*

```
%sum: $B:XBEB:GXX1:BG1:Z&zi#be§-zm:BZ:zi#zm:BZ:zi#zm:BZ:zi#zm
$0:XBEB:GTS1:CB1:M&T-T#crawl§-across-path-to-right:MZ:zi#zm:MZ:zi#zm:MZ:zi#zm
$R:XBEB:GXX1:RF1:M&T-T#crawl§-onto-path-onto-side:RZ:zi#zm:RZ:zi#zm:RZ:zi#zm
```

*SUIJ: *a little mouse [c] sneaking, over to the table [c] and climbed up the table [c], got the cheese [c] and slid down [c] and went down to his mouse hole [c] to eat the cheese [c].*

```
%sum: $B:XSOU:GXX2:BG1:Z&zi#be§-zm:BG1:M&T-T#sneak§-over-to-table:BZ:zi#zm:BZ:zi#zm
$M:XSOU:GTS1:CB1:M&T#climb§-up-table:MZ:zi#zm:MZ:zi#zm:MZ:zi#zm
$D:XSOU:GTS1:CB1:M&T#slide§-down:MZ:zi#zm:MZ:zi#zm:MZ:zi#zm
$R:XSOU:GXX1:RF1:Z&T-T#go§-down-to-mouse-hole:RZ:zi#zm:RZ:zi#zm:RZ:zi#zm
```

*SUIJ: *Il fait son jogging [c] et il passe sur la route [c] et il traverse jusqu'à l'autre bout en courant*

```
%sum: $0:XCOU:GLS4:MR1:M&zi#faire-jogging§-zm:MR1:T$#passer§-sur-route:CR2:T&T-
M#traverser§-jusqu'à-l'autre-bout-en-courant:MZ:zi#zm
$B: XCOU:GXX0:BZ:zi#zm:BZ:zi#zm:BZ:zi#zm:BZ:zi#zm
$R: XCOU:GXX0:RZ:zi#zm:RZ:zi#zm:RZ:zi#zm:RZ:zi#zm
```


Sub-field – Vendler Typology: Predicates are classified in four types depending on their intrinsic properties in accordance with two dimensions – the existence of a right boundary and durativity:

AST ‘State’
 ACT ‘Activity’
 ACC ‘Accomplishment’
 ACH ‘Achievement’
 ZZZ (absence of verb)
 Delimiter = hash (#)

Sub-field – Durativity

DUR durative
 DUN non-durative
 DNC durativity uncodable
 DZZ (absence of verb)
 Delimiter = hash (#)

Sub-field – Boundedness

BOY Bounded
 BON Unbounded
 BNC Boundedness uncodable
 BZZ (absence of verb)

After this fourth sub-field, ‘§-’ signs are inserted, followed by information relating to the predicate boundary (verb or periphery):

- z: no boundary or boundary unknown.
- vmm, vmp, vcm: boundary expressed by a motion verb comprising a right boundary: (*traverser, poser, faire entrer*).
- ep: boundary expressed by a preposition: *across, into, jusqu’à, au bout de, zu, rein*’.

Examples:

S0#AST#DUR#BON§-z: *there’s a nest in a tree, c’est Popi, da ist eine Raupe*
 S1#ACT#DUR#BON§-z: *he looks at the nest il pousse le ballon, er rennt auf der Straße, Hoppy is pushing a suitcase down a small hill, Popi tire une voiture de course.*
 S2#ACC#DUN#BOY§-vmm: *Hoppy is pushing his case down to the trail, Popi traverse la rue du village, Popi entre dans la grotte.*
 S2#ACH#DUN#BOY§-vmm: *The cat stopped, ‘Un chat arrive.’*
 S2#ACH#DUN#BOY§-ep: *He climbs up to the top, il monte jusqu’en haut de la dune.*
 SZ#AZZ#DZZ#BZZ§-z: *no verb Hoppy on the beach, Popi dans la montagne, Popi im Wald.*

Field 5: Verb type

The verb list adopts some codes from the %cod lines, and adds new codes that enable more detailed analysis:

VMM#| motion verb (*walk, bounce, traverser, monter, entrer, laufen, rennen*).
 VMP#| caused motion verb (*push, pull, transitive roll, pousser, schieben, ziehen*).
 VCM#| manner verb in explicit causative construction (*faire rouler*).
 VCX#| causative construction with dynamic verb other than a motion verb (*faire appeler*).

VSS#	static verb (<i>be, stand, être, avoir, sein, liegen</i>).
VPC#	causative construction with caused motion verb (<i>faire pousser</i>).
VXX#	other dynamic verbs (<i>eat, catch, get, boire, licken</i>).
VSZ#	static verb absent.
VMZ#	motion verb absent.
VPZ#	caused motion verb absent.
VXZ#	other dynamic verb absent.
VZZ#	absent verb that cannot be inferred from the context.

Additional codes:

VPP#	perception verb (<i>see, hear, voir, apercevoir, hören</i>)
VCD#	cognitive / modal verbs (<i>think, want, réfléchir, mögen</i>)
VPPZ#	absent perception verb.
VGZ#	absent cognitive modal verb.

After #: semantic information as on %cod lines: +Tbo +Tve +Tdi +Tdx +Tso +Tgo + Tpa +Mma +Mmp + Cse + Mac

Additional codes marking temporal information are added once the quotation form of the verb is added after the ‘|’bar sign, and separated from the quotation form by the §-:

+Bor	for coding the verb or all parts of the predicate that indicate a right temporal boundary
+Bow	for coding the verb or all parts of the predicate that can indicate a right temporal boundary but where the encoding in the language is ambiguous
+Ite	for coding of the verb and other means that express the sense of iteration: <i>again</i>
+Stp	Storypast as in : <i>once in once upon a time</i>
+Dur	Durativity: <i>still</i> in the example <i>He's going up, he's still going up</i>
+Sim	Simultaneity: <i>at the same time</i>

Example: *the bird returns with an earthworm in its beak*
 = S2#ACH#DUN#BOY§-vmm:VMM#+Tdi |return§-Ite

Field 6: Verbal morphology: inflection (see %cod lines)

Field 7: Breakdown of verbal inflections: Time (different from %cod lines)

TPA	past: <i>the dog chased the cat</i>
TNP	not past – present: <i>the dog catches the cat by the tail</i>
TAU	others: infinitive, gerund: <i>Le chat ayant peur du chien</i>
TIN	indeterminate
TZZ	no inflection

Field 8: Breakdown of verbal inflections: Aspect (additional field)

AIMP	imperfect: <i>La chat ayant peur du chien</i>
APE	perfect: <i>A bird made its nest on the branch of a tree</i>
ANM	unmarked: <i>et on voit le chat au pied de l'arbre</i>
ANP	non progressive (in the case of an available marked aspectual pair in verbal morphology)
APR	progressive (in the case of an available marked aspectual pair in verbal morphology)
AIN	non codable (including ambiguous): <i>the bird # return [= ? returned] home [/] # to his home # with a # squirrel [/]</i>
AZZ	no inflection

Fields 9 to 13: Temporal and spatial information outside of the verb

- EP – marker linked to space –prepositional expressions
 EL – marker linked to space –particles
 EX – marker linked to space – types (other than EP and EL)
 EO – marker linked to neither space nor time, but relevant
 EZ – absence of marker

Additional codes:

- ET – marker linked to temporality (including adverbials, connectives)
 CNT – other markers: *Et le chat s'en va ...*

After the '&' sign, semantic information is indicated.

- For ET and CNT: +Ite, + Stp, + Dur, + Imm, +Pun, + Ant, + Post, + Sim, +Oth
 For EP, EL, EX and EO: See %cod lines.

- | | |
|--|-----------------------------------|
| - <i>Hoppy goes up the hill again</i> | = ET1/1&+Ite again |
| - <i>He is still pushing the suitcase</i> | = ET1/1+Dur still |
| - <i>So once upon a time there was a little family of birds</i> | = CNT1/3&+Stp so+once-upon-a-time |
| - <i>He stopped for a while to look at the nest</i> | = ET1/1&+Dur for-a-while |
| - <i>Hoppy goes all the way to the top of the hill</i>
way-to-top-of+NP | = EP1/1&+Tbo+Bor+Loc all-the- |
| - <i>Before he was at the top of the hill</i> | = ET1/1&+Oth before |
| - <i>Popi tire la chaise jusqu'au garage</i> | = EP1/1&+Tbo+Bor jusqu'à+NP |
| - <i>Popi va dans la grotte</i> | = EP1/1&+Loc+Bow dans+NP |
| - <i>Mais à ce moment là un chien arrive</i> | = CNT1/2&+Oth mais-à-ce-moment-là |
| - <i>D'un seul coup la maman oiseau s'en va</i> | = ET1/1&+Pun d'un-seul-coup |
| - <i>Ensuite on voit la mère qui s'envole</i> | = ET1/1+Post ensuite |
| - <i>Pendant ce temps-là le chat lui mord la queue</i> | = ET1/1+Sim pendant-ce-temps-là |

NB. When two or more connectives immediately follow each other in the utterance, they are coded in the same CNT field (example 1); otherwise, they are coded separately (example 2).

- Example 1: *And then the cat runs away*: CNT1/1&+Oth|and-then
 - Example 2: *Et le chat s'enfuit donc*: CNT1/2&+Oth|et:CNT2/2&+Oth|donc

Field 14: Periphery: independent main (matrix) clauses and subordinate clauses.

Coding of the periphery is the same as in the %cod line, with the following exception: For subordinate clauses (PS), occurrences of the below are indicated in a sub-field:

+Sim, +Bor, +But

- +Sim – Simultaneity (Example: *while, quand*)
 +Bor – Temporal perspective responding to the question 'When?' implying a right temporal boundary, but one that is not necessarily achieved (Example: *jusqu'à, atteindre, pénétrer*)
 +But – Causal perspective responding to the question 'Why?' and implying the individual's intentions, but not necessarily implying that the goal is achieved (Example: *in order to catch the babybirds, pour aller dans la maison*)

- *The cat licks its lips **looking** at the nest and the little ones:* PS+Sim|0+ppr
- *He pushes the gift **while** he goes up the hill:* PS+Sim|when+pr
- *Il monte à l'arbre **pour dévorer** les petits:* PS+But|pour+inf
- *Il pousse le cadeau **jusqu'à arriver** en haut de la colline:* PS+Bor|jusqu'à+inf

Field 15: Figure (see %cod lines)

Field 16: Ground (see %cod lines)

Field 17: Comments in brackets { } (see %cod lines)

Important remarks:

The coding of predicates includes the verb and all its arguments. Certain spatial expressions used with certain Grounds can give rise to confusion with regard to the temporal boundary they express. It is advisable to indicate sources of ambiguity wherever they arise by using a similar code for all fields, i.e., via the code 'Bow' in the %tem line, and the 'W' code in the %sut line (cf. below).

Locative cases where boundary marking is ambiguous: *go in, get on, aller dans, aller sur*. All these spatial expressions are best coded as ambiguous both on the %tem and %sut line.

%tem line: ...EPx/n&+Loc+Bow|

Example: *Popi va dans la grotte avec son caddie*

%tem: ... SW#A WW#DUR#BNC§-z:VMM#|aller:PR:TNP:ANM:EP1/2&+Loc+Bow|dans+NP

Expressions of the type *tout en haut, tout en bas*, etc. that are coded +Loc on the %cod lines must be coded +Loc+Bor on the %tem and LB on the %sut lines, as they not just indicate spatial information but also mark an temporo-aspectual right boundary:

%tem: ...:EPx/n&+Loc+Bor|tout-en-haut§-emp:

Expressions of the type *jusqu'en haut de la montagne* consequently are coded as follows:

%tem: ... EPx/n&+Tbo+Loc+Bor|jusqu'en-haut-de+NP

Idiosyncratic expressions:

Some of the verbal and other expressions indicating boundaries are idiosyncratic, especially in the L2 learner data. For example: *au croisé de, il bougeait la table entre la ferme*. In those cases, code: +Oth+Loc+Bow on the %tem line.

%tem: ...EPx/n&+Oth+Loc+Bow| au-croisé-de+NP§-idiolex-AmbVprep-across/a-traversé

%tem: ...EP1/1&+Oth+Loc+Bow |entre+NP§-idiolex-AmbVprep(=into)

%sut: ...LW...

Synthetic lines for temporality: %sut

These lines have the same number of fields and the same syntax globally as the synthetic lines for space. However, they also indicate the temporal boundaries for each clause with letters (B, N, and W) which are added immediately after the semantic content to which they relate:

B – bounded

N – unbounded

W – ambiguous with respect to boundedness

The below example indicates (by means of the ‘N’ following the ‘CB’ coding) that the clause is unbounded:

- *Il monte la dune en poussant...*

CBN2:T&CA#il-monte-la-dune§-en-poussant:...

2.4 Suggestions for the coding of L2 learner and patient data

Adult learners of a second or third language as well as aphasic patients are more likely than children or adult native speakers to have word-finding problems, to produce idiosyncratic utterances or turns of phrases, and to produce errors that make the coding of their data less straightforward. Adult L2 learners are also likely to use lexical items from their native language when they fail to find an appropriate expression in their L2. To deal with such issues, and with the resulting productions of our L2 learners and aphasic patients, we have provided some additional coding and other tips below.

2.4.1 Utterances corresponding to a ‘lexical negotiation’ (NGL)

Some utterances in L2 learner and aphasic patient data (possible also in child data, but much less frequently) are used in order to discuss a word or expression with the experimenter (often in cases of lexical gaps in the participants’ vocabulary). In these cases, the code NGL (‘lexical negotiation’) is shown in field 2 of the coding line for this clause, and the rest of the utterance is not coded. When the negotiation concerns the verb or satellites, do *not* code NGL, so as not to lose relevant information for the analyses, but code normally.

A code often co-occurring with the NGL code in Field 2, is the code GBE (Given By Experimenter) in the comment field. In the below example, the first line produced by an L2 learner is left uncoded but marked with NGL, but the second line is coded as normal but with an indication of the word given by the experimenter.

*SUI: He's pulling the..... I don't know what it's called?
 \$:NGL:0:0:0#0|0:0:0:0&0|0:0&0|0:0&0|0:0&0|0:0:Fi0|0:Gr0|0:{ngl}
 *EXP: pram.
 *SUI: He's pulling the pram.
 \$:CIB0:.....: {ngl+poussette-GBE}

2.4.2 Code switching (cs)

Another possible phenomenon in adult L2 data (and bilinguals) is the occurrence of code switching. Code-switching can be defined as the practice of alternating between two or more languages or varieties of language in conversation. It causes issues for the coder, as some of the expressions in the utterance of relevance for the coding, i.e., verbs, satellites may be in a language other than the target language, and the information contained may therefore be slightly different from information that can typically be expressed in this language. In the case of code switching, use the code -cs.

The code should be placed in the field concerned (e.g., in field 5 if it relates to the verb, in fields 8-12 if it concerns the satellites, etc.). General comments are also indicated in field 16, for example, in the case where the subject goes from one language to another for a whole section of their discourse. In this case, we use the symbol cs with the same syntax as for auto-correction, e.g. {cs-words1a-2a-3a/words1b-2b-3b}.

Some specific idiosyncrasies may occur in the context of code switching, specifically when a learner assumes that a verb in their language can be found as a direct translation in the L2 as well (they assume there to be a cognate). For example, our English learners of French would sometimes assume that the verbs *ascendre* (=go up), *pouller* (=pull), *crosser* (=cross) are legitimate verbs in French, whereas in fact they are not. In these cases, where the meaning of those verbs can easily be understood, code as follows (indicating that the expression is a result of code-switching, and is idiosyncratic (idiov)).

- *Popi ascend le bouée sur la maison* = VMP#+Cse+Tve|ascend§-cs-**idiov**-monter-trans:

2.4.3 Categorial ambiguities (verb/satellite)⁵

In the particular L2 data that we were working with (English learners of French) we found instances in which the L2 learners used expressions that were based on verb stems in the target language, but were satellites in the source language. Similar examples could also occur in the reverse learning direction (French learners of English). In these cases, it will often be difficult for the coder to decide if the element should be coded as a verb or as a satellite, hence the title of this section. A code was devised specially for these situations:

AmbVprep: Ambiguity of syntactic function of an item (verb or preposition).

The code 'AmbVprep' can be used for ambiguous expressions between verbs and satellites in field 5 (verb type) as well as in fields 8-12 (satellites). They are at least of three types:

- a) If the clause contains just one problematic element that can be interpreted as a verb, it is coded in field 5. NB: AmbVprep is followed by a hyphen, then the inferred word.
- *he acrossed the street*
Field 5 = VMM#+Tbo|across§-AmbVprep-cross
Field 8 = EZ
- b) On the other hand, if the clause contains a clear verb and another ambiguous kind of element (i.e., ambiguous between verb and preposition), the problematic element is coded in fields 8-12.
- *Il poussé la roue entre la cave*
Field 5 = VMP for *pousser*
Field 8 = AmbVprep for *entre* EP1/1&+Tbo|entre+NP§-AmbVprep(=into)
- c) Lastly, some cases introduce ambiguity that is more difficult to resolve between the use of an element as a verb or as a satellite. In these cases, and in the absence of any indication that enables the presence of a verb to be inferred (for example, the inflection in *he acrossed the street*, see above), the form is coded as a satellite in fields 8-12 by indicating the omission of the verb in field 5.
- *he across the street*
%cod: \$:CIB0:LR01:SX:VZZ#:FZ:DOT:EP1/1&+Tbo|across+NP§-cs-AmbVprep:EZ2/1&|:EZ3/1&|:EZ4/1&|:EZ5/1&|:PM0:Fig|PRO(popi):Grd|street: { }

2.4.4 Idiosyncratic usages

For idiosyncratic usages in L2 learners and aphasic patients, the coding provides the coder with the freedom to identify these in the respective fields as well as in the commentary. Some examples of such codes are provided below. Note that the same codes can be used for the native speakers (child and adult), but should be less likely to occur in those populations. As a general rule, code whatever can be inferred from the utterance in the case of idiosyncratic use of expressions:

§-idioep:

Idiosyncratic choice of a spatial prepositional expression.

- The child go with one side of the road on the other.
EO...+Tso|with...§-idioep-from... = EP...+Tgo|on...§-idioep-to
- *Popi pousse la valise dans le toit...* = EP1/1&+Loc|dans+NP§-idioep
- *Il tire la malle sur l'épaule.* = EP1/1&+Loc|sur-épaule§-idioep

§-idiom:

Idiosyncratic morphology

- He brought the bag up the hill = VMP#+Cse|bring§-idiom-brought

⁵ Note that categorial ambiguities may also occur in child language data, but are less frequent overall

- *Popi a descend la colline de neige avec le ballon* = VMM#+Tve|descendre§-idiom-a-descend

§-idioaux:

Idiosyncratic use of an auxiliary

- *Popi est roulé le roul de le voiture entre la cave* = VMP#+Cse+Mmp|rouler§-trans-**idioaux**-être

§-repasp:

Repetition in order to provide temporal-aspectual value (durativity/iterativity)

- What's Hoppy doing? Rolling, rolling, rolling =VMP#+Cse+Mmp|roll§-repasp
 - The caterpillar is going up, up, up =EL#+Tve|up§-repasp
 - *Il monte, il monte, il monte.* = VMM#+Tve|monter§-repasp
 - *Er schiebt und schiebt bis der Tisch da ist* = VMP#+Cse+Mac|schieben§-repasp

§-idiops: Idiosyncratic choice of a subordination marker.

- *Hoppy # to pull the box [c] when he comes down the hill [c]. PS/when§-idiops*

§-Idiopm0: Absence of main phrase.

- *Popi [c], montant la dune [c]* = PS/en+ger§- idiops

§-idionoprep: A preposition is compulsory in a structure, but is not used.

- *le caddy entre [dans] le ferme.* VMM#+Tbo|entrer§-intrans-idionoprep(entre-ferme)

§-idioeo: Idiosyncratic choice of a relevant non-spatial expression.

- *Papy est traversé la rue # et il tirer avec une petite cheveau*
 EO1/1&|avec- cheveau§-idioeo-idiolox
He ran fastly = idioeo

§-idiocontract: Idiosyncratic contraction or absence thereof.

- *Popi tire le sac jusque en haut de la colline*
 EP1/1&+Tgo|jusque-en-haut-de+NP§-idiocontract

Below we provide all coding annotations for idiosyncratic uses of the target language in alphabetical order, as they can be used for the various types of data sets that a project of this type might be engaged with.

Table 8: Coding annotation

CODE	EXPLANATION	EXAMPLE	CODING
AmbVprep (field 5)	Ambiguity of syntactic function of an item (verb or preposition)	<i>Popi est roulé le roul de le voiture entre la cave</i>	§:CIB0:...VMP#+Cse+Mmp rouler§- trans- idioaux(être)...:EP1/1&+Tbo entre+NP§-idiolox- AmbVprep(=into):... Fig le-roul-de- le-voiture§-roue-
cs	Cases of code switching	<i>Il pulled le sac, il a pulled le sac</i>	VMP... pull§-cs
Idioaux (field 5)	Idiosyncratic auxiliary	<i>il est monté la valise</i>	VMP#&+Cse+Tve monter§- idioaux
idiocase	Apparent or possible error of case	<i>Er überquert der Strasse.</i>	Grd straße§- idiocase

idiocontract	Contraction or absence of idiosyncratic contraction	<i>Popi tire le sac jusque en haut de la colline</i>	EP1/1&+Tgo jusque-en-haut-de+NP§- idiocontract :
idioeo	Idiosyncratic choice of a relevant non spatial expression (fields 8-12)	‘Hoppy crossed the road # and he pulled with a little horse’	EO1/1& with(com)+NP§- idioeo -idiolex:
idioep	Idiosyncratic choice of a spatial prepositional expression (fields 8-12)	‘Hoppy is pushing the suitcase in the roof ’	EP1/1&+Loc in+NP§- idioep :
idiofig	Idiosyncratic choice of Figure	<i>il, tiré, le devant de caddy</i>	Fig le-devant-de-caddy§-idiofig- idiofig
idiog	Apparent or possible error of gender	<i>Popi monte la voiture sur la toit</i>	Grd toit§- idiog
idiogrd	Idiosyncratic choice of Ground (Field 14-15)	<i>he had the gift [c] push it on the house [c]</i>	Grd house§- idiogrd (= on the roof of the house)
idioimp	Figure or Ground implicit but idiosyncratic in the context of the clause and/or the item	<i>it's on a house [c] he rolls [the rubber ring] [c] going up(wards)</i>	Fig imp(rubber-ring)§- idioimp
idiolex	Idiosyncratic lexis and sometimes invented	<i>Popi pousse la roue dans une caverné dans les bois</i>	Grd une-caverné+les-bois§- idiolex
idiom	Idiosyncratic morphology (Field 5)	<i>Popi a descend la colline de neige avec le ballon</i>	VMP#&+Cse+Tve descendre§- idiom (a-descend)
idionoaux	Obligatory auxiliary absent	<i>he going up</i>	VMM# go§- idionoaux
idionodet	Absence of a nominal determinant, which is obligatory in the context of the phrase	<i>Popi monte voiture.</i> ‘Hoppy go up car’	Fgr voiture§- idionodet
idionoprep	Omission of an obligatory preposition (Field 5)	<i>Il marche la maison</i>	VMM#&+Mma marcher§- idionoprep
idionov	Idiosyncratic absence of verb	<i>Elle... trottinette dans la rue.</i>	VMZ# §-trottinette-idionov VMZ# §-up- idionov

idiopho	Indicates the case where the learner uses a form that is idiosyncratic on the phonetic level. These indications enable an understanding of the reason a coder has chosen to code one linguistic form rather than another.	[<i>tene</i>]	traîneur§-idiopho
idiops	Idiosyncratic choice of a subordinate marker or the absence of a main clause (Field 13)	<i>Hoppy pulls the box [c] when he comes down the hill [c].</i>	PS when§-idiops:
idiov	Frequent idiosyncratic verbs in the case of code switching or ‘calque’ = <i>ascendre, pouller, crosser</i>	<i>Popi ascend la bouée sur la maison</i>	ascend§-trans-cs-idio(v)(monter)
trans/-intrans	For each verb that can be used transitively and intransitively, indicate if it is used in a context where semantic transitivity is or is not marked (Field 5)	<i>Il roule la boule en haut de la colline.</i> <i>Parce qu’elle roule.</i>	rouler§-trans rouler§-intrans

2.5 Information relevant to the coding of Distractor items

2.5.1 Introductory remarks

Some of our experimental designs include distractor items. Typically, distractor items do not get coded or analysed, but occasionally one may find that the distractor items produce linguistically interesting phenomena that may be worth analysing after all. Our Hoppy distractor items are of this type. We therefore designed, in a later phase, a coding specifically for these items which all included events that happened as a result of a non-agentive (inanimate) object moving into other objects which then moved as a result (*falling over; sliding along the floor, bouncing into a wall, etc.*). For the full list of distractor items, see Appendix x in Volume 2.

In the below we will describe the only field that is different from the %cod lines.

2.5.2 Field 2: Item Identification

Similar to the %cod lines for the experimental items, Field 2 indicates if the coded clause is expressing the main motion information (CIB0), information regarding to the setting (BEG) or information regarding to the resulting state. Coding for CIB0 and resulting state clauses have been altered on the distractor coding lines, as described below.

Clauses referring to the Beginning (no change compared to %cod lines)

An example of a clause coded BEG is given below. This coding is not different from BEG on %cod lines in the

experimental items.

Well there were two equal sized balls a blue one and a red one (BEG)

CIB0 Coding lines

For the purposes of the distractor items, however, the main motion information (CIB0) is further coded, depending on what type of information is expressed. Thus, all CIB0 components are systematically further divided into four sub-components, identified by their individual sub-codes in this field. The sub-codes are separated from the CIB0 code by the §- delimiter.

The components correspond to the following types of information:

CIB0§-FM indicates that the following codes concern the motion of the object that is the instantiating Force of the complete sequence of Motion events hence evolving, as exemplified in the example below. Note that if two FM motion events are expressed, a sequential number after 'FM' is used, such that FM1, FM2 as in the second example below:

- *The blue ball rolled forward (CIB0§-FM)*
- *it's rolling on a flat surface (CIB0§-FM1) going from right to left (CIB0§-FM2)*

CIB0§-FX, in contrast, marks that the following codes concern the instantiation by the Force object of the motion event through a non-motion verb, as in the following 2 examples:

- *La boule jaune a donné de l'élan au livre (CIB0§-FX)*
- *La boule jaune est lancée de gauche à droite (CIB0§-FX)*

CIB0§-IPC indicates the expression of the Impact of the object as in the example below, i.e., it indicates that a motion event led to an impact into another object.

CIB0§-FM-IMP: when this impact is expressed in the same clause with the expression of the instantiating motion of the Force, the two codes are combined as in CIB0§-FM-IMP:

- *and hit a red ball (CIB0§-IPC)*
- *The red ball rolls into the dice (CIB0§-FM-IMP)*

The codes CM and/or CX, following CIB0§- indicate the expression of a causing event (whether it causes further motion (CM) or another result (CX)) as in the following examples:

- *there's a blue ball (BEG) # **pushing the red ball** (CIB0§-CM)*
- *qui vient frapper un vase (CIB0§-IPC) et **qui le fait tomber** (CIB0§-CM) heu, **en le brisant.** (CIB0§-CX)*

CIB0§- followed by QOM, QOX, QOS and/or QOC identify clauses expressing the consequence of the motion of the instantiating Force on the Object or of the causing event, either through the use of a motion verb (QOM), another type of action verb (QOX), a static verb (QOS) or a caused motion expression.

- *and the book's **going back and forth.** (CIB0§-QOM)*
- *knocked over the wine bottle [c] which **is now broken.** (CIB0§-QOX)*
- *so all the dice **are on the ground.** (CIB0§-QOS)*

- *le livre rebondit sur le mur (CIB0§-QOM) et retourne jusqu' à la boule jaune (CIB0§-QOM) pour la toucher (CIB0§-QOX) et la fait pas bouger (CIB0§-QOC).*
- *Right, there's now a yellow ball [BEG] and it's rolled [CIB§-FM1] and pushed a red book [CIB0§-CM] which has slid across [CIB0§-Q1OM] bounced off the black wall [CIB0§-Q2OM] and rebounded to the yellow ball [CIB0§-Q3OM].*

Clauses referring to the end of a sequence of events in a given item

Rather than coding all clauses referring to the end of the sequence of events projected in the video clip, in the distractor items only the final motion of either the instantiating Force object or the final motion / resulting state of the Object are coded, as in the examples.

- *the blue ball stopped. (FSEF)*
- *and there it stopped. (FSEO)*

An example of the various codes used in the second field is provided below for one of the distractor items:

Alors la boule verte (BEG) qui roule de gauche à droite (CIB0§-FM1), et qui vient frapper un vase (CIB0§-IPC) et qui le fait tomber (CIB0§-CM1) heu, en le brisant (CIB0§-CX2) et qui s'arrête immédiatement (CIB0§-FSEF).

Coding of all other fields are as on the %cod lines.

3. SPONTANEOUS PRODUCTIONS

For coding spontaneous productions (for example: longitudinal data from young children), there are two possible options: a traditional system, and the new one described above.

3.1 Traditional system

A previous coding system exists for young children's spontaneous productions⁶. This system is much simpler than the one described above for the experimental studies, but it is less accurate, and insufficient for certain types of analysis. Briefly, this system comprises 12 fields organised according to the following syntax:

```
%cod:(tab)$S:situation:verb|xx:infl:expression-spa1|xx:expression-spa2|xx:expression-spa3|xx:otherexpressions|xx:conn|xx:figure|xx:ground|xx:{comment}
```

1. Introduction of the coding line (\$S).
2. Situation type (static or dynamic).
3. Verb type, followed by the verb infinitive.
4. Verbal inflection (see codes).
5. 1st preposition or other spatial expression: type, numbering, form used.
6. 2nd preposition or other spatial expression (ditto).
7. 3rd preposition or other spatial expression (ditto).
8. 'Non-spatial' preposition or adverbial expression: type, followed by form.
9. Type of connective between utterances, followed by the form used.
10. Located entity, followed by the entity explicitly or implicitly referred to.
11. Entity in relation to which the Figure is located, followed by entity explicitly or implicitly referred to.
12. Coder comments in brackets { }.

3.2 Adaptation of the new system

For analyses specifically devoted to the representation of space in a cross-linguistic perspective, the most satisfactory option consists of adapting the system described in section II to spontaneous productions, which, moreover, enables a direct comparison of this type of data with that gathered in more controlled experimental situations. As a reminder, the fields of the new system are the following:

1. Introduction of the line (\$).
2. Script component and identification of the target.
3. Response type (spontaneous or induced, numbering).
4. Situation type (static, dynamic or indeterminate).
5. Verb type and information expressed by the verb|verb infinitive.
6. Verbal inflection (varies according to the language, coding is adjusted for specific subjects).
7. Voice.
8. 1st satellite (includes prepositions, particles, adverbs, etc.) & information expressed by the satellite|form.
9. 2nd satellite (ditto).
10. 3rd satellite (ditto).
11. 4th satellite (ditto).
12. 5th satellite (ditto).
13. Periphery: independent matrices and subordinate clauses.
14. Figure: indication of the referent functioning as the Figure.
15. Ground: indication of the referent functioning as the Ground.
16. Comments in brackets { }.

In this second option, field 2 (useful for analysing targets in experimental data) is no longer relevant and can be

⁶ Hickmann, M. & Champaud, C. (2007). *L'émergence de l'espace dans la langue de l'enfant: manuel de codage*. Laboratoire Structures Formelles du Langage, CNRS & Université de Paris 8.

replaced by four letters beginning with the upper case 'C' (for example, CXXX), or can be used to code a new dimension of data which could be useful for analysing spontaneous production – on condition that certain constraints linked to Chiles and the programmes aimed at analysis are respected (NB. The letter 'C' is only necessary if the programme specifically written for analysis is used).

Moreover, field 3, corresponding to the type of response, can be adapted to take account of interaction (questions from the child, responses by the child to a question from the adult, etc.).