Systemic Functional glossing conventions

Conventions for interlinear unit-by-unit glosses for lexicogrammar.

Interlinear unit-by-unit glosses give information about the functions, classes and lexical translations for linguistic items (potentially of any extent). To date Systemic Functional linguists have used a variety of different types of notation and glossing across publications. The different information and level of detail provided often makes it difficult to compare descriptions across languages and contrast alternative descriptions of the same language; it also risks producing misunderstandings and difficulties in comprehension (see Lehmann 1982 for similar concerns about morphemic glossing outside SFL). Nonetheless, there are many conventions that are shared across the research community, but these are not necessarily used in each paper. The main purpose of this document is to make the most widely used conventions explicit and help standardize Systemic Functional glossing. Some innovations are suggested, however these are relatively minor and mostly optional.

The glossing conventions detailed here are intended for descriptions informed by Systemic Functional Linguistics (hereafter SFL), especially those involving tiers of function structures. They are designed to complement rather than replace the 'Leipzig' morpheme-by-morpheme glossing rules and can be deployed in conjunction with the Leipzig rules. This is especially important for analyses including morphemic glossing and/or papers with an intended audience beyond the SFL community. As these conventions complement the Leipzig rules, a number of them adopt the strategy of generalizing the Leipzig conventions from morphemic glossing to the glossing of other units. If the analysis requires only a morphemic glossing, without specifying function structures, the Leipzig rules should be used. The conventions developed here are however written with an understanding that grammatical distinctions are regularly distributed across multiple ranks and a single tier of morphemic glossing is not explicit about this distribution. The Leipzig glossing rules can be found here:

https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf

An author's arguments and readers' needs will affect the level of detail required. Accordingly, like the Leipzig rules, the conventions specified here allow some flexibility with various options noted. In the spirit of SFL, the conventions should be interpreted as a set of resources for glossing, rather than a list of rules.

The conventions have been written in a general form so they can be used across all lexicogrammatical units. For the Spanish rank scale (Quiroz 2013), for example, this means they can be generalized for clauses, group/phrases, words and morphemes. Note here that these conventions are not system network conventions. System networks have their own conventions and formalisms which are introduced in detail in Martin et al. (2013). The glossing conventions specified here do not attempt to account for phonological, discourse semantic, register or genre analysis. To deal with these strata further development of these conventions is required.

In line with the Leipzig rules, it is important to note that there are often multiple ways of describing language patterns. Glossing is thus a form of analysis, rather than part of the data. The resources here

aim to offer standardized means of showing glosses in line-by-line analyses. Each convention is illustrated with examples that gloss different units. Many examples are also given in conjunction with the Leipzig conventions, as these are r
equired by many journals and publishers. The glossing of many examples below has been modified for the purposes of this document and so the references in each example should be read as 'adapted from'. Examples without references have been supplied by researchers, as listed at the end of this document.

The reference to cite these conventions is:

> Systemic Language Modelling Network (2018) *Systemic Functional glossing conventions*. https://systemiclanguagemodelling.com/glossing/

Convention 1A: Basic organization of the glossing, unit-by-unit alignment and translation

Each gloss should include the example in romanized script using appropriate punctuation and italics, a glossing of class and/or function (see Convention 2), a unit-by-unit translation and a translation of the full example. Interlinear glosses are left-aligned vertically, unit by unit, with the example. If analyzing groups within a clause, for example, each group will be left-aligned vertically with its gloss. Classes are given in lower case and may be abbreviated or specified in full where appropriate. Translations of each unit are also left-aligned vertically below the class or function analysis. This is most effectively done by using the table function in word processors.

romanized script	*Tengo*	*dos codificadores*	*en mi casa.*
class and/or function	verbal group	nominal group	prepositional phrase
unit-by-unit translation	*I have*	*two set-top boxes*	*in my house*
full example translation	'I have two set-top boxes in my house.'		

Below the glossing, a translation of the whole example should be given in single quotation marks '…'. This translation may be an 'idiomatic' translation of the example or it may attempt to show more clearly the organization of the object language example itself. This will depend on the needs of the discussion (Rose 2005). Examples using class glossings are:

(1) Chilean Spanish group analysis within clause

Tengo	*dos codificadores*	*en mi casa.*
verbal group	nominal group	prepositional phrase
I have	*two set-top boxes*	*in my house*
'I have two set-top boxes in my house.'		

(2) Korean word analysis within nominal group

hyeoksinjeokin	*cheotbeonjjae*	*jeongi*	*jadongcha*
adjective	numeral	noun	noun
innovative	*first*	*electric*	*car*
'innovative first electric car'			

(3) Gija group analysis within clause

tany lalangkarrany	*warrk nginini*
nominal group	verbal group
that crocodile	*he.danced*

'That crocodile danced.'

Convention 1B: Writing system alignment (optional)

If showing the original writing system is useful, it may be included as the top line. If the writing system unfolds left-to-right and is segmentable in line with the romanized transliteration, the appropriate segments may be left-aligned with other glossings.

(4) Korean group analysis with Leipzig morphemic glossing

오고	있나	보다
o-go	*it-na*	*bo-da*
verbal group		
come-LNK	be-LNK	seem-NDEFER;DECL;FML
come	*be ...ing*	*it.seems*

'seems to be coming'

If the writing system has a different orientation (right to left, top to bottom etc.) or cannot be easily aligned in this way, but it is still included, then it should be kept in its original format without any alignment.

(5) Mongolian group analysis in clause

ᠪᠢ ᠡᠷᠯᠦᠭᠡ ᠢᠷᠡᠵᠦ ᠳᠡᠶᠢᠯᠬᠦ ᠦᠭᠡᠢ ᠃

bi	*örlöge*	*irejü deilxü ügei*
nom.gp	nom.gp	verb.gp
I	*morning*	*unable to come*

'I am unable to come in the morning.'

Convention 1C: Leipzig morphemic glossing and unit-by-unit translation (optional)

If morphemic analysis is needed in addition to the analysis of other units, the Leipzig morphemic glossing may replace the unit-by-unit translation, or both may be used. In this case, the romanized example should use the Leipzig rules regarding hyphens for segmentable morphemes, equals signs for clitics etc.

(6) Korean (Shin 2018) group analysis in clause plus morphemic analysis

    a.

        saram i   *man-ket-da*
        nom.gp   verb.gp
        people NOM   there.be.many-would-NDEFER;DECL;FML
        'There would be many people.'

    b.

        saram i   *man-ket-da*
        nom.gp   verb.gp
        people NOM   there.be.many-would-NDEFER;DECL;FML
        people   *there would be many*
        'There would be many people.'

(7) Gija group analysis in clause plus morphemic analysis

    ta-ny   lalangkarra-ny   warrk   ngini-ni
    nominal group   verbal group
    that-M   crocodile-M   dance   3SG.M-do.PST
    that crocodile   *he.danced*
    'That crocodile danced.'

Convention 1D: One-to-many correspondences

If the translation of a single unit necessarily involves the use of a higher level unit or multiple units of the same type, these may be separated by a period. This convention is a generalization of Leipzig Rule 4. Such glossing is only needed if it is important to make explicit the two different sets of units being used. For example, in Chilean Spanish the word *conmigo* is translated by the English phrase *with me*, and so periods are used between words in the English translation.

(8) Chilean Spanish word level glossing

    Bea   *fue*   *conmigo*   *al*   *cine*
    proper noun   verb   pronoun   preposition noun
    Bea   *went*   *with.me*   *to*   *movies*
    'Bea went with me to the movies.'

Or in Gija, the main participant is often marked through a prefix on the verb, in a way that cannot be translated by a single word in English:

(9) Gija word level glossing

warrk	*nginini*
coverb	verb
dance	*he.did*

'he danced'

Convention 2: Function structure glossing

Often, grammatical functions will be needed for the analysis. Functions should take initial capital letters (e.g. Actor), and should be left-aligned. Functions may occur in place of, or in addition to, a class analysis.

(10)   Chilean Spanish (Quiroz 2008: 49) nominal group function structure

a.

los	*canales*
Deictic	Thing
the	*channels*

'the channels'

Where functions and classes are both specified, there are two alternatives. One involves inserting the function above the class that realizes it:

b.

los	*canales*
Deictic	Thing
det.	c.noun
the	*channels*

'the channels'

Alternatively, functions and the classes realising them may be placed on the same line. Here, the function comes first followed by an arrow → and then the class:

c.

los	*canales*
Deictic→det.	Thing→c.noun
the	*channels*

'the channels'

Convention 3A: Conflation of functions or classes

Conflation of functions on a higher rank realized by a single unit on the rank below are marked by a forward slash /. Slashes may also be used for an object-language element that is formally unsegmentable but involves the conflation of two or more classes (i.e. the co-selection of two features leading to portmanteau realizations). Note that this relates to Leipzig rule 4B concerning formally unsegmentable elements with distinguishable grammatical properties, where a semi-colon ; is used rather than a slash /. For class glossing, especially morphemic classes, the Leipzig semi-colon is an acceptable alternative. Conflated functions however should always use a forward slash /.

(11) Chilean Spanish (Quiroz 2008: 49) verbal group function structure

no	*los*	*cambia*
Neg	P-clitic	Finite/Event
not	*them*	*it.changes*

'It doesn't change them.'

(12) Tagalog clause classes

Hindi uminom ang babae ng beer.
negative/declarative/material/medium-focus
'The woman didn't drink beer.'

Convention 3B: (Optional)
An alternative to using a forward slash / for conflating functions is to place functions realising the same unit on different lines. This is especially useful where there is not a one-to-one correspondence between different functions, often involving different metafunctions. This convention can also be used for expansion of functions (see Convention 10).

(13) French (Caffarel 2006:171) clause functions

Hier,	*le capitaine*	*est arrive.*
Theme	Rheme	
Circumstance	Actor	Process
Yesterday	*the captain*	*arrived*

'Yesterday, the captain arrived.'

If more explicitness is needed about the boundaries of functions or classes, or they are not entirely clear from the spacing, table borders may be used. In the following example, table borders are used to show the boundaries of each function, which makes explicit the full extent of the Rheme.

(14) French (Caffarel 2006:171) clause functions

Hier,	*le capitaine*	*est arrive.*
Theme	Rheme	
Circumstance	Actor	Process
Yesterday	*the captain*	*arrived*

'Yesterday, the captain arrived.'

Convention 4: Multi-level glossing

Sometimes, relevant grammatical information is distributed across ranks and so multiple levels of glossing is useful, either for functions or for classes or both. In this case, an additional line of glossing may be added. The glossing should be left-aligned at each level, with the higher levels having larger spaces between elements and the lower levels having smaller spaces. As with Convention 3B, borders may be used to make explicit the boundaries at each level. It is optional whether a translation should be given for each level. Examples 15a., b., c. and d. show different ways in which one or more layers of glossing can be included.

(15)   Korean clause, group and word analysis

a.

seonnyeo	*sen*	*man*	*naeryeowatda*
P1→nom.gp			Process→verb.gp
only three nymphs			*descended*
Thing→noun	Quant→num.	IFM→part.	Event→verb
nymphs	*three*	*only*	*descended*

'Only three nymphs descended'

b.

seonnyeo	*sen*	*man*	*naeryeowatda*
P1→nom.gp			Process→verb.gp
Thing→noun	Quant→num.	IFM→part.	Event→verb
nymphs	*three*	*only*	*descended*

'Only three nymphs descended'

c.

seonnyeo	*sen*	*man*	*naeryeowatda*
P1			Process
nominal group			verbal group
Thing	Quantifier	IFM	Event
noun	numeral	particle	verb
nymphs	*three*	*only*	*descended*

'Only three nymphs descended'

d.

seonnyeo	*sen*	*man*	*naeryeowatda*
P1			Process
nominal group			verbal group
only three nymphs			*descended*
Thing	Quantifier	IFM	Event
noun	numeral	particle	verb
nymphs	*three*	*only*	*descended*

'Only three nymphs descended'

As many publications require the use of the Leipzig glossing rules, these can be followed for morphemic glossing.

(16) Korean clause and morpheme analysis

eonni	*geureom*	*i*	*saeg*	*eun*	*eotta-e*	*yo*
Theme					Rheme	
Sister	then	this	colour	TOP	how.is-INFML	DEFERENCE

'Sister, well then, how is this colour?'

(17) Brazilian Portuguese clause and morpheme analysis

cê	*cont-a*	*seu*	*sonho*	*pra ele*
Subject	Predicator	Complement		Adjunct
nom. gp	verb. gp	nom. gp		prep. phrase
2SG	tell-3SG.IMP	2SG.POSS.M dream		to 3SG.M

'Would you tell him your dream?'

Convention 5: Complexing

Complexing occurs when two elements from the same rank occur together to realize a function at a higher rank. The simplest means of glossing complexes is to specify the class as a unit complex in the glossing:

(18)   Pitjantjatjara (Rose 2001: 436) clause analysis with group complexing

wati tjuta minyma tjuta munu tjitji tjuta	*tjungurinkula*
Medium	Process
nominal group complex	verbal group
men, women and children	*coming together*

'men, women and children are coming together'

If it is useful to specify the elements realizing a complexed unit, this may be glossed as below. Note that the hierarchy of realization here involves a unit complex realizing a function at the higher rank (e.g. Medium→nominal group complex); unit complexes are then realized by complexing functions (e.g. nominal group complex→1$^+$2$^+$3); complexing functions are then realized by simplex classes (e.g. 1→ nominal group). Linkers such as *muna* in the following example often do not occur within the simplex class being complexed so should be glossed through a translation or a lower level glossing.

(19)   Pitjantjatjara (Rose 2001: 436) clause analysis with group complexing

wati tjuta	*minyma tjuta*	*munu*	*tjitji tjuta*	*tjungurinkula*
Medium				Process
nominal group complex				verbal group
1	+2		+3	
nom.gp	nom.gp		nom.gp	
men	*women*	*and*	*children*	*coming together*

'men, women and children are coming together'

(20) Mandarin Chinese clause analysis with group complexing

shenjing xitong	*jieshou*	*zonghe*	*he*	*chuandi*	*gefangmian xinxi*
Actor	Process				Goal
nom. group	verbal group complex				nom. group
	+1	+2		+3	
	verbal.gp	verbal.gp		verbal.gp	
the nervous system	*receives*	*synthesises*	*and*	*transmits*	*all aspects of information*

'The nervous system receives, synthesises, and transmits all aspects of information.'

Convention 6A: Embedding

Embedding (downranking) occurs when a unit from one rank occurs in a unit of the same or lower rank. Embedded clauses are marked by double square brackets [[…]], and embedded groups/phrases are marked by single square brackets [..] in the glossing. The brackets should be left- and right-aligned with the beginning and end of the elements being embedded. Elements that are downranked to another rank should keep their normal labelling. For example if a clause is embedded in a group, it should be labelled as a clause, but placed in the same line as the group glossing.

(21) Mongolian embedded clause in a nominal group

bi irejü deilxü ügei		*örlöge*
Qualifier		Thing
[[clause	]]	noun
I am unable to come		*morning*

'the morning that I am unable to come'

(22) Korean embedded co-verbal phrase within a nominal group

minjujueui e daehan		*siheom*	*munje*
Qualifier		Classifier	Thing
[co-verbal phrase	]	noun	noun
about democracy		*examination*	*question*

'examination question about democracy'

Convention 6B: Further analysis of embedding (optional)

If it is useful to analyse the structure of embedded elements, this analysis should continue downward following conventions. As in Convention 3B, table borders may be used to explicitly mark boundaries of classes or function.

(23) Korean embedded co-verbal phrase within a nominal group

minjujueui	*e*	*daehan*	*siheom*	*munje*
Qualifier			Classifier	Thing
[co-verbal phrase		]	noun	noun
Incumbent		Role		
[nom.gp		] bound verb		
Thing	EFM			
noun	particle			
democracy		*about*	*examination*	*question*

'examination question about democracy'

Convention 6C: Layering (nesting)

When a complex element is layered within another complex element, the layered elements are surrounded by parentheses (…).

(24) Pitjantjatjara (Rose 2001: 401) clause complex layering

ka watjanu	*meeting panya runamilantjaku*	*ngananya nyura Chairman tjungkuntjikitja mukuringanyi*	
1	2(β	α	)
and said	*in order to run this meeting*	*who do you3 want to put as Chairman?*	

'and said in order to run this meeting who do you want to put as Chairman?'

As complexing often produces very long stretches of language, an alternative convention is to arrange the complexes vertically, with the function labels for each layer. Nesting with complexes can be shown through vertical alignment; note the layout below for $\alpha \,^{x}\beta \,^{x}\gamma$ (1 $^{+}$2).

(25)   Pitjantjantjara (Rose 2001: 418) clause complex layering

α *tjana startara communitynya moneylta ngalyungu*
        'they started to give the money to the communities

$^{x}\beta$ *tjarantjaku*
        to divide it up

$^{x}\gamma$ 1   *anangu tjutangku community runamilantjaku*
        so the people could run the communities

        $^{+}$2   *or homelandakutu ankuntjaku*
        or move out to their homelands'

Convention 7: Non-overt elements and ellipsis

Implicit wording, where acknowledged, may be marked by ø in the object-language; alternatively the wording can be made explicit and placed in parentheses. In either case glosses should be placed in parentheses. Implicit meanings should only be recognized where an explicit alternative realization has been formalized the description. Analysts must be careful not to impute a class, function or more detailed glossing for implicit elements without significant argumentation.

(26) Pitjantjatjara (Rose 2004: 513) clause analysis

kukua kanyilaku	ø	*tatinu*	*pulingka*
Purpose	(Actor)	Process	Place
for wallaby game	*(they)*	*climbed up*	*in the hills*

'For wallabies, that is, they climbed up in the hills.'

(27) Tagalog (Martin and Cruz 2018) clause analysis

pa-gamit	(*ako*)	*ng kalan*
Process	(Actor)	Scope
	(Theme)	
FA.cause.NFIN-use	(T.1SG)	NTP stove

'Let me use the stove.'

Convention 8: Discontinuous elements

Elements that are discontinuously realized across two or more positions in a unit can be indicated by '…' – eg. Function…Function or class…class. This notation can be used for prosodic structures or circumscribing syntagms (e.g. circumfixes). As with previous conventions, borders may be used to specify the boundaries of classes or functions if needed.

(28) Tagalog (Martin and Cruz 2018) clause analysis

kailangan	*pa*	*ba*	*–ng*	*tulung-an*	*kita*
Modul	Terms…			Predicator	…Terms
need	IMM	Q	–LK	help-FA.NFIN	NTP.1SG/T.2SG

'Do I still need to help you?'

(29) Chilean Spanish group analysis

cuatro	*amigos*	*más*
Numer...	Thing	…ative
word...		…complex
α		β
numeral	common noun	adverb
four	*friends*	*more*

'four friends more'

Note that discontinuous realization signaled by … is not to be confused with multiple instantiations of the same function or class. Multiple instances of the same function or class simply repeat the labelling.

(30) Tagalog (Martin 2004: 273) clause analysis

Nagtapos	*ng AB in Economics, summa cum laude*	*sa De La Salle University*	*noong 1972.*
Process	Goods	Location	Location
completed	*his BA in Economics, summa cum laude*	*at De La Salle University*	*in 1972*

'(He) completed his BA in Economics, summa cum laude at De La Salle Univerity in 1972.'

Convention 9: Interrupting or interpolating elements

Interrupting or interpolating elements are enclosed by angled brackets <<…>> in the gloss.

(31) Korean projected clause interrupting projecting clause

그는	"소년들의 건강상태가 양호합니다"	라고	말했다.
geu neun	*"sonyeon-deul ui geongangsangtae ga yanghoha-mnida"*	*rago*	*malha-et-da.*
1…	<<"2>>		…1
he	*"the boys' health condition is good"*	*LK*	*said*

'He said "the boys' health condition is good". '

Convention 10: Expansion of functions

When a function is expanded into two or more functions (such as Negotiator expanding to Subject, Finite, Predicator and Clitics in French as in Caffarel (2006), or Mood expanding to Subject, Finite in English as in Halliday and Matthiessen (2014)), the expanding functions is to be written on a line below the expanded function.

(32)   Tagalog (Martin and Cruz 2018) clause functions and morphemic glossing
a.

ma-sarap	*daw*		*ang pagkain*
Negotiator			Scope
Predicator	Terms		
STATIVE-tasty	REPORTATIVE	T	food

'They say the food is tasty'

Convention 11: Multiple levels of delicacy

If it is useful to specify a class to multiple levels of delicacy, a colon : is used.

(33) Tagalog (Martin 1996: 264) clause classes
ibinayad ng babae ang 200 pesos
clause:material:explosive
'The woman paid the 200 pesos.'

References

Caffarel, A. (2006) *A Systemic Functional Grammar of French: From Grammar to Discourse.* London: Continuum.

Halliday, M. A. K. and Matthiessen, C. M. I. M. (2014) *Halliday's Introduction to Functional Grammar*. London: Routledge.

Lehmann, C. (1982) Directions for interlinear morphemic translation. *Folia Linguistica*. 16: 199-224.

Martin, J. R. (1996) Transitivity in Tagalog: A functional interpretation of case. In M. Berry, C. Butler, R. Fawcett and G. Huang (eds) *Meaning and Form: Systemic Functional Interpretations*. New Jersey: Ablex. 229-296.

Martin, J. R. (2004) Metafunctional profile of the grammar of Tagalog. In A. Caffarel, J. R. Martin and C. M. I. M. Matthiessen (eds) *Language Typology: A Functional Perspective*. Amsterdam: John Benjamins. 255-304.

Martin, J. R., Wang Pin and Zhu Yongsheng (2013) *Systemic Functional Grammar: A Next Step into the Theory: Axial Relations*. Beijing: Higher Education Press.

Martin, J. R. and Cruz, P. 2018 Interpersonal grammar of Tagalog: a systemic functional perspective. *Functions of Language,* Special Issue on 'Interpersonal meaning; Systemic Functional Linguistics perspectives' edited by J R Martin. 25:1. 54-97.

Quiroz, B. (2008) Towards a systemic profile of the Spanish MOOD. *Linguistics and the Human Sciences*. 4:1. 31-65.

Quiroz, B. (2013) *The Interpersonal and Experiential Grammar of Chilean Spanish: Towards a Principled Systemic-Functional Description Based on Axial Argumentation.* Unpublished PhD Thesis. Department of Linguistics, The University of Sydney.

Rose, D. (2001) *The Western Desert Code: An Australian Cryptogrammar*. Canberra: Pacific Linguistics.

Rose, D. (2004) Metafunctional profile of the grammar of Pitjantjatjara. In A. Caffarel, J. R. Martin and C. M. I. M. Matthiessen (eds) *Language Typology: A Functional Perspective.* Amsterdam: John Benjamins. 479-536.

Rose, D. (2005) Narrative and origins of discourse: Construing experience in stories around the world. *Australian Review of Applied Linguistics*. Series S(19). 151-177.

Shin, G (2018) Interpersonal grammar of Korean. *Functions of Language*. 25:1. 20-53.

Unless otherwise noted, examples without references were supplied by:

Korean: Gi-Hyun Shin, J. R. Martin; *Mongolian*: Dongbing Zhang; *Brazillian Portuguese*: Giacomo Figueredo; *Chilean Spanish*: J. R. Martin, Beatriz Quiroz; *Gija*: Anna Crane; *Tagalog*: J. R. Martin; *Mandarin Chinese*: Hao Jing, Wang Pin.

These conventions were developed collaboratively by members of the Systemic Language Modelling Network (SLAM). In particular, they were developed based upon a series of talks and workshops given by Giacomo Figueredo, Wang Pin, Y. J. Doran and J. R. Martin. However the development of these conventions was truly collaborative involving detailed feedback and suggestions both written and face-to-face by a number of people, including Dongbing Zhang, Gi-Hyun Shin, Beatriz Quiroz, Hao Jing, Anna Crane, Harni Kartika Ningsih, Mira Kim, Andres Ramirez, Estela Moyano, Alice Caffarel, David Rose, Lungguh Ariang Bangga, Laura Hlavacka, Teresa Oteiza, Margarita Vidal, Tom Bartlett, Lise Fontaine, Ayako Ochi and Mohamed Ali Bardi.