

ICHL Oxford, 2022 August 2

# Is cliticization an intermediate stage between free lexeme and affix status?

MARTIN HASPELMATH

*Max Planck Institute for Evolutionary Anthropology (Leipzig)*

Overview:

- what we understand about grammaticalization is that **bound grammatical** markers often come from **free content words**, and that their ordering is more **rigid** – but not much else (§1)
- the idea of a **continuous scale** (synchronic and diachronic) from lexemes to affixes via clitics is a widespread stereotype, but it has not been tested systematically (§2-3)
- we need **clear (and simple) definitions** of the terms *lexeme*, *clitic*, and *affix* (§4)
- clitics differ from affixes primarily in that they are **indiscriminate** (= do not exhibit word-class selectivity), not in that they are “**less phonologically attached**”
- phonological “attachment” is not clearly defined and does not clearly distinguish clitics from affixes
- most types of **clitics do not become affixes**; and most types of **affixes did not pass through an intermediate clitic stage**
- clitics **may** be an intermediate stage in affix development, but this is **unusual**

## 1. What we understand about grammaticalization

- grammatical markers often have their origin in a content word or a demonstrative
- content words and demonstratives are typically **free** forms (occurring on their own), while grammatical markers are (almost by definition) **bound** forms
- three types of changes:
  - **free form > bound form**
  - **long form > short form**
  - **shortening leads to rigidification of order**

(the longer a form/expression, the more flexible its order)

What we don't understand:

- whether “analytic” languages tend to become “synthetic” languages  
(cf. Haspelmath & Michaelis 2017)
- whether “agglutinative” languages tend to become “fusional”  
(cf. Haspelmath 2017)
- whether morphology and syntax are different components of grammar  
(cf. Haspelmath 2011a)

## 2. A stereotype: lexeme > clitic > affix

Typical textbook accounts of grammaticalization claim that cliticization is an intermediate stage between free lexeme and affix status (e.g. Hopper & Traugott 2003: 7; Wiemer 2014: 442; Narrog & Heine 2021: 279; Konvička 2019).

(9) content item > grammatical word > clitic > inflectional affix (> Ø)

Spencer & Luís (2012: §2.1): “Inflections are generally thought to arise in historical language change through a stage of cliticization”

For example, free personal pronouns are said to become clitic pronouns first before they turn into person-number affixes on verbs (“Givón’s Cycle”, Seržant 2021).

schematically:

*Jane, she comes.*  
 > *Jane (,) she=comes.*  
 > *Jane she-comes.*

The processes of cliticization and affixation are sometimes collapsed as “coalescence” (e.g. Haspelmath 2011b), suggesting that they constitute a single **gradual** process.

## 3. A continuous scale?

Is there a continuous scale (or “cline”) from lexicality to affixhood?

(or even discourse > syntax > morphology > zero; Givón 1979)

This may make sense at an intuitive level, but in order to test it scientifically, we need ways of **measuring an item’s position** on the scale. Rigorous quantitative testing requires clear qualitative comparative concepts.

So far, no attempt has been made at quantifying grammaticalization changes, as far as I know.

Few attempts at defining “affix” and “clitic” have been made (Zwicky & Pullum 1983 provide a set of “diagnostic symptoms”, which are not really criteria; cf. Haspelmath 2015).

The problems with defining *word* are well-known, too (Haspelmath 2011a; Tallman 2020). Clearly, a wide range of different phenomena are involved in “**tightness of bonding**” or “**coalescence**”, and multiple dimensions need to be taken into account (Börjars et al. 2008). But these dimensions need not correlate (van Gijn & Zúñiga 2014).

## 4. Defining affixes and clitics

For the great majority of cases, the following definitions give results that correspond to linguists' intuitions about the meanings of these terms:

(1) **affix** (Haspelmath 2021)

An affix is a bound morph that is not a root and that always occurs on roots of the same class (nouns, verbs, adjectives; Haspelmath 2022a)

(2) **clitic** (Haspelmath 2022b)

A clitic is a bound morph that is neither a root nor an affix.

A **free form** differs from a **bound form** (a clitic or an affix) in that it can **occur in isolation** (e.g. Zwicky & Pullum 1983: 502).

- It is clear that when a free form (e.g. a noun or a verb or a demonstrative) grammaticalizes, it finally becomes a **bound form** – “grammatical status” (Boye & Harder 2012) is much the same as “bound status”.

e.g.	Old Church Slavonic	<i>vlast ta</i>	‘that power’
		<i>ta</i>	‘that one’
	> Bulgarian	<i>vlast-ta</i>	‘the power’
		<i>*ta</i>	‘that one’ (tazi)

- It is also clear that grammaticalization involves shortening. Grammaticalized forms are **bound** and **short**.

- But does a free form generally become a **clitic** before it becomes an affix? (Apparently not.)

## 5. Clitics differ from affixes in being indiscriminate (“promiscuous”)

### 5.1. Focusing clitics

(3) Russian

- a. *Pročita-la li Anna knigu?*  
 read-PST PQ Anna book  
 ‘Did Anna read a book?’

- b. *Knigu li Anna pročita-la?*  
 book PQ Anna read-PST  
 ‘Did Anna read a BOOK?’

- c. *Včera li Anna čita-la?*  
 yesterday PQ Anna read-PST  
 ‘Did Anna read YESTERDAY?’

## (4) Lak (Kazenin 2002: 293)

- a. *Na q̄atri d-ullali-ṣa =ra.*  
 I house(G4) G4-build.DUR-PTCP=1SG  
 ‘I am building a house.’
- b. *Na =ra q̄atri d-ullali-ṣa.*  
 I =1SG house(G4) G4-build.DUR-PTCP  
 ‘The one who is building a house is me.’
- c. *Na q̄atri =ra d-ullali-ṣa.*  
 I house(G4) =1SG G4-build.DUR-PTCP  
 ‘What I am building is a house.’

## 5.2. 2nd-position clitics

## (5) Tagalog (Kaufman 2010: 10)

- a. *Na-túto =siya nang= wika =ng Intsik.*  
 AV-learn =3SG.NOM GEN= language =LNK Chinese  
 ‘She learned Chinese.’
- b. *Hindí =siya na-túto nang= wika =ng Intsik.*  
 NEG =3SG.NOM AV-learn GEN= language =LNK Chinese  
 ‘She didn’t learn Chinese.’

## (6) Serbo-Croatian (Bošković 2001: 8)

- a. *Zašto smo mu je predstavili juče?*  
 why 1PL.SBJ him.DAT her.ACC introduced yesterday  
 ‘Why did we introduce her to him yesterday?’
- b. *Predstavili smo mu je juče.*  
 why 1PL.SBJ him.DAT her.ACC yesterday  
 ‘We introduced her to him yesterday.’

## 5.3. Epiphrasal clitics (= clitics occurring at the periphery of a phrase)

## (7) English (postphrasal)

- a. [*my friend*]*’s house*  
 b. [*the lady I met yesterday*]*’s offer*  
 c. [*the boy I like*]*’s new bike*

## (8) German (prephrasal)

- a. *nur* [*mein Hund*]  
 b. *nur* [*in ihrem Haus*]  
 c. *sie will nur* [*spielen*]

## 5.4. Other kinds of clitics

## (9) Colloquial German

- a. *Willst de se haben?*  
 want.2SG you them have  
 ‘Would you like to have them?’ (Standard: *Willst du sie haben?*)

- b. *Wenn de se haben willst...*  
 if you them have want.2SG  
 ‘If you want to have them...’ (Standard: *Wenn du sie haben willst...*)

### 5.5. “Phonological attachment”?

It is sometimes thought that clitics differ from affixes in that they are “phonologically attached” to a lesser degree. But is this the case? What does “attachment” mean?

## 6. Clitics are usually unstressed, and may be welded (like affixes)

Clitics are usually **unstressed** (in languages with stress). In fact, they show even less of a tendency to be stressed than affixes, because they tend to occur outside the stress domain.

- e.g. Turkish *gel-di-Ø* ‘she came’  
*gel-dí mi?* ‘did she come?’  
*gel-di-níz* ‘you (PL) came’

- e.g. Russian *krug* [kruk] ‘circle’  
*krúg li?* [krukli] ‘circle?’  
*krug-i* ‘circles’

In these cases, affixes do not look like “reduced variants” of clitics.

Clitics, not only affixes, may show segmental interaction with adjacent words (they may be **welded**; Haspelmath 2021):

- (12) Turkish (polar question particle)  
*geldí mi?* ‘did she come?’  
*öldü mü?* ‘did he die?’  
*Alí mi?* ‘Ali?’  
*dün mü?* ‘yesterday?’

- (13) Russian (prepositions *v(o)* ‘in’, *s(o)* ‘with’)  
*v nužde* ‘in need’ *vo vrede* ‘in harm’ (\**v vrede*)  
*s radost’ju* ‘with joy’ *so straxom* ‘with fear’ (\**s straxom*)

- (14) English (indefinite article)  
*a tree* *an old tree*

Clitics are probably less commonly welded than affixes, but at least in principle, **segmental phonological interaction is independent of “clitic vs. affix” status**.

More generally, we need a better definition of the informal notion of “phonological attachment”.

## 7. Most types of clitics do not become affixes

- **discourse particles** rarely become affixes

- (15) a. *However*, our ambitious proposal failed.  
 b. Our ambitious proposal, *however*, failed.  
 c. Our ambitious proposal failed, *however*.

- **2nd position clitics** rarely become affixes

- (16) Ancient Greek (Lowe 2016: 377)  
*Ho dè Kûros ekselaúnei epì tòn potamón.*  
 DEF but Cyrus march.army to DEF river  
 ‘But Cyrus marches his army to the river.’

- **coordinator clitics** rarely become affixes

- (17) Latin  
 a. *ménsa* ‘the table’  
 b. *mensá=que* ‘and the table’

- **question particles** rarely become affixes

- (18) Turkish  
*geldi mi?* ‘did she come?’  
*öldü mü?* ‘did he die?’  
*Ali mi?* ‘Ali?’  
*diin mü?* ‘yesterday?’

### However:

It is true that person indexes do seem to become affixes in **some** cases, e.g. in the Romance languages. Old Portuguese and Old Spanish seem to attest a **clitic stage** of the weak person forms (Pescarini 2021).

- (19) Old Portuguese  
*logo lhe el-rrei taxava que ...*  
 soon 3.DAT the king ordain.IPFV that  
 ‘immediately the king imposed on him as punishment that...’

- (20) Old Spanish  
*Sy el físico la bien connosçe*  
 if the physician 3.F.ACC well know.3SG  
 ‘if the physician knows it well’

Modern Romance languages do not allow this sort of “interpolation”: Object clitics must **immediately precede** the verb, or **immediately follow** the verb (i.e. they have affixal status).

In Polish, the development from 2nd position clitics to verbal person markers seems to be a very typical case of gradual grammaticalization.

Andersen (1987): *Nigdy=śmy tego nie myśleli.* >  
*Nigdy tego nie myśleli-śmy.* ‘We never thought that.’

(5)	Total no. of examples	Deviations from Wackernagel's rule		Agglutination to preterite form	
1500's	580	12	2%	130	23%
1600's	1303	64	4%	649	49%
1700's	1439	62	4%	994	68%
1800's	1988	308	15%	1395	80%
1900's	3325	503	15%	2817	84%
exp. prose	569	usual		525	92%

## 8. Most types of affixes did not pass through a clitic stage

The most typical affixes:

**verbs:** tense-aspect and person affixes  
**nouns:** case, number and gender affixes

Most of these do not derive from elements that are usually called clitics.

The grammaticalization origin of **number** and **gender** affixes is often unclear, though **articles** seem to play an important role. Articles may be clitics, but they may also be affixes (= occur always on the noun).

There do not seem to be clear cases of affixal articles deriving from clitic articles

(the Nordic suffixed article does not derive from a clitic; Börjars & Harries 2008)  
 e.g. *\*barn hit > barn-it > barn-et* ‘the child’

Tense-aspect affixes from **auxiliary verbs** (e.g. Bybee & Dahl 1989),  
 but auxiliaries are very often non-indiscriminate (class-selective)

e.g. Spanish *mis amigos han llamado*  
 my friends have called

Case affixes are generally thought to derive from **auxiliary (relational) nouns** (e.g. König 2011),  
 but such “auxiliary nouns” are very often non-indiscriminate

e.g. Lezgian *t'ara-n k'an-ik* (Haspelmath 1993: 220)  
 tree-GEN bottom-LOC  
 ‘under the tree’ (see also Belyaev 2010 on Ossetic)

But it is true that **person affixes** may sometimes go back to promiscuous person forms (see above for Romance and Polish).

## 9. Conclusion

- Many linguists have the intuition that there are degrees of bondedness – more “**tightly bonded**” grammatical forms and more “**loosely bonded**” forms
- Often linguists seem to assume that the tightness of bonding (or “coalescence”) is of a phonological nature (they often talk about “**phonological boundness**”) – but there is no clear conceptual basis for this
- Comparative grammar research needs clear comparative concepts, and *clitic* can be defined clearly only with respect to their **indiscriminate** behaviour (contrasting with class-selective affixes)
- It is of course plausible that grammaticalization involves not only **shortening** (= fewer segments) and **boundness** (= non-occurrence in isolation), but also some kind of “phonologically close association”, but this has not been formulated clearly yet.
- At present, the well-known scale “**lexeme > clitic > affix**” is not more than a stereotype, and we need to replace it by a more rigorous formulation.

## References

- Andersen, Henning. 1987. From auxiliary to desinence. In Harris, Martin & Ramat, Paolo (eds.), *Historical development of auxiliaries*, 21–52. Berlin: Mouton de Gruyter.
- Belyaev, Oleg. 2010. Evolution of case in Ossetic. *Iran and the Caucasus* 14(2). 287–322. (doi:10.1163/157338410X12743419190269)
- Börjars, Kersti & Harries, Pauline. 2008. The clitic-affix distinction, historical change, and Scandinavian bound definiteness marking. *Journal of Germanic Linguistics* 20(4). 289–350. (doi:10.1017/S1470542708000068)
- Boye, Kasper & Harder, Peter. 2012. A usage-based theory of grammatical status and grammaticalization. *Language* 88(1). 1–44.
- Bybee, Joan L. & Dahl, Östen. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13(1). 51–103.
- Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.
- Haspelmath, Martin. 2011a. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1). 31–80. (doi:10.1515/flin-2017-1005)
- Haspelmath, Martin. 2011b. The gradual coalescence into “words” in grammaticalization. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford handbook of grammaticalization*, 342–355. Oxford: Oxford University Press.
- Haspelmath, Martin. 2015. Defining vs. diagnosing linguistic categories: A case study of clitic phenomena. In Błaszczak, Joanna & Klimek-Jankowska, Dorota & Migdalski, Krzysztof (eds.), *How categorical are categories? New approaches to the old questions of noun, verb, and adjective*, 273–304. Berlin: De Gruyter Mouton.
- Haspelmath, Martin. 2021. Bound forms, welded forms, and affixes: Basic concepts for morphological comparison. *Voprosy Jazykoznanija* 2021(1). 7–28. (doi:10.31857/0373-658X.2021.1.7-28)
- Haspelmath, Martin. 2022a. Word class universals and language-particular analysis. In van Lier, Eva (ed.), *Oxford handbook of word classes*. Oxford: Oxford University Press (to appear).
- Haspelmath, Martin. 2022b. Types of clitics in the world’s languages (to appear).
- Hopper, Paul J. & Traugott, Elizabeth Closs. 2003. *Grammaticalization: 2nd edition*. Cambridge: Cambridge University Press.
- [https://www.academia.edu/83478708/Types\\_of\\_clitics\\_in\\_the\\_worlds\\_languages](https://www.academia.edu/83478708/Types_of_clitics_in_the_worlds_languages)



- Kaufman, Daniel. 2010. *The morphosyntax of Tagalog clitics: a typologically driven approach*. Ithaca: Cornell University.
- Kazenin, Konstantin I. 2002. Focus in Daghestanian and word order typology. *Linguistic Typology* 6(3). 289–316. (doi:10.1515/lity.2003.002)
- König, Christa. 2011. The grammaticalization of adpositions and case marking. In Narrog, Heiko & Heine, Bernd (eds.), *The Oxford handbook of grammaticalization*. Oxford: Oxford University Press.
- Konvička, Martin. 2019. Grammaticalisation clines: A brief conceptual history. *History and Philosophy of the Language Sciences* 2019. (<https://hiphilangsci.net/2019/03/06/grammaticalisation-clines/>)
- Lowe, John J. 2016. Clitics: Separating syntax and prosody. *Journal of Linguistics* 52(2). 375–419. (doi:10.1017/S002222671500002X)
- Monachesi, Paola. 1999. *A lexical approach to Italian cliticization*. Stanford: CSLI Publications.
- Narrog, Heiko & Heine, Bernd. 2021. *Grammaticalization*. Oxford: Oxford University Press.
- Pescarini, Diego. 2021. *Romance object clitics: Microvariation and linguistic change*. Oxford: Oxford University Press.
- Samvelian, Pollet & Tseng, Jesse. 2010. Persian Object Clitics and the Syntax-Morphology Interface. 212–232. (<https://web.stanford.edu/group/cslipublications/cslipublications/HPSG/2010/abstr-samvelian-tseng.shtml>)
- Seržant, Ilja A. 2021. Cyclic changes in verbal person-number indexes are unlikely. *Folia Linguistica* 55(s42–s1). 49–86. (doi:10.1515/flin-2021-2014)
- Spencer, Andrew & Luís, Ana R. 2012. *Clitics*. Cambridge: Cambridge University Press.
- Tallman, Adam J. R. 2020. Beyond grammatical and phonological words. *Language and Linguistics Compass* 14(2). e12364. (doi:10.1111/lnc3.12364)
- Wiemer, Björn. 2014. Quo vadis grammaticalization theory? Why complex language change is like words. *Folia Linguistica* 48(2). 425–468. (doi:10.1515/flin.2014.015)
- Zwicky, Arnold M. & Pullum, Geoffrey K. 1983. Cliticization vs. inflection: English n't. *Language* 59(3). 502–513.
- Zwicky, Arnold M. 1994. What is a clitic? In Nevis, Joel A. & Joseph, Brian D. & Wanner, Dieter & Zwicky, Arnold M. (eds.), *Clitics: a comprehensive bibliography 1892-1991*, xii–xx. Amsterdam: Benjamins.