(Anti-)cliticization in Mapudungun

Fernando Zúñiga

Abstract This article describes the formal behavior of some elements found in Mapudungun (isolate, South America): a question particle, a postposition, and several 3rd-person markers. Framed in terms of current theories of phonological and grammatical words, the paper argues that a useful characterization of the Mapudungun elements under scrutiny should acknowledge (a) that clitics are interestingly heterogeneous regarding how different bound elements stand in paradigmatic opposition to each other, and (b) that some of these elements can be meaningfully be called anti-clitics (i.e., they are p-words that are part of larger g-words).

Keywords Clitic · Anti-clitic · p-word · g-word · Mapudungun

1 Introduction[1]

Mapudungun is an unclassified language spoken by around 250,000 people in south-central Chile and west-central Argentina. Fluency rates seem to be declining in younger generations, albeit not too sharply; the highest concentrations and/or numbers of fluent speakers are found among (elderly) adults living in rural, or at least non-metropolitan, areas in the Bío-Bío and Araucanía Regions (Chile). The present study concentrates on the variety called Central Mapudungun, mainly spoken in Araucanía. Unless otherwise specified, data come from my own field notes. The orthographic convention used here for rendering Mapudungun data is the so-called Alfabeto Mapuche Unificado.

---

[1] I am indebted to Rik van Gijn for numerous and valuable comments to a previous version of this article.

F. Zúñiga (✉)
Department of Linguistics, University of Bern, Länggassstrasse 49, 3012 Bern, Switzerland
e-mail: fernando.zuniga@isw.unibe.ch

The morphological-typological profile of Mapudungun is not rare in the Amerindian context. With respect to synthesis, the verbal domain is polysynthetic, the nominal domain is mildly synthetic, and other domains are analytical. (Here, polysynthesis means productive nominal incorporation, systematic verb stem serialization, and a relatively large number of potential affixal slots on verbs.) With respect to the technique parameters, the largely agglutinative morphology can be more precisely classified as concatenative, of predominantly separative exponence (i.e., polyexponential morphemes and portmanteaus are rare), and without flexivity (i.e., allomorphy is found in some inflectional phenomena, but it is almost exclusively phonologically conditioned). Basically suffixing and (clause-)head-marking, Mapudungun shows constituent order patterns that reflect both syntactic and pragmatic sensitivity: typical clauses are of the V(S) type, especially in narrative texts, but VS, VO and SVO clauses are also found, the latter type especially in elicitation.

The stress rules of Mapudungun are important for the present paper, so the basics are outlined in what follows. Disyllables in isolation are stressed depending on syllable structure (with closed syllables attracting stress), e.g. *rúka* 'house', *máñke* 'condor', *pewé<u>n</u>* 'monkey puzzle tree', and *tralkán* 'thunder'. Monomorphemic trisyllables have secondary stress on the antepenult, e.g. *àchawáll* 'hen'. Plurisyllabic verb forms are usually somewhat more complex, with some inflectional suffixes attracting stress, and words in actual phrases behave according to more intricate patterns that are not well understood yet. See Zúñiga (2006:63f) for more details.[2]

Armed with a simple and reasonably mainstream typology of clitics (Sect. 2), the present paper analyzes three different kinds of elements that appear appended to other elements in Mapudungun and have received comparatively little attention from a form-oriented perspective in the literature. (In fact, neither cliticization in particular nor wordhood in general are topics treated in great detail in Mapudungun studies.) Section 3 deals with the question particle *am*, Sect. 4 addresses adpositions, paying special attention to the postposition *mew*, and Sect. 5 presents different 3rd person markers that show interesting properties against the background of any theory of wordhood. Section 6 summarizes the findings and gives some conclusions. The paper argues that, like with most natural-language phenomena that are approached theoretically and typologically with a simple but powerful analytical toolkit, it is the intermediate and hard-to-capture cases that are particularly interesting. For Mapudungun, this amounts to acknowledging that not all clitics, or their mirror images ("anti-clitics"), are equal, and that characterizing precisely such elements in more detail leads to a better understanding of how the morphosyntax of this Amerindian language, as well as morphosyntax in general, really works.

---

[2]My own observations coincide with the ones presented originally by Adalberto Salas in Salas (1978) and later in Salas (2006:73–75). These are largely compatible with Smeets (2008:49–50), who also makes crucial reference to the *right* edge of the phonological word, but incompatible with Echeverría and Contreras (1965) (the sole source for e.g. Goedemans and Van der Hulst 2013), who describe a stress pattern centering on the *left* edge of the word.

2 Typologies of clitics

Most widespread notions of clitic revolve around the notion of autonomy. Prototypical words are autonomous both syntactically and phonologically, while prototypical affixes are dependent, i.e. non-autonomous, in both respects. Many current scholars see prototypical clitics as syntactically autonomous but phonologically dependent (i.e., they are g-words in their own right but part of a larger p-word, cf. Aikhenvald 2002).

A number of subtypes of such intermediate elements have been proposed in the literature. One of the most popular taxonomies distinguishes between simple and special clitics, based on whether such dependent elements behave according to the default syntactic rules of the language or not (cf. Halpern 1998; Zwicky 1977, 1985; Zwicky and Pullum 1983; but see Anderson 2005 for a critique of simple clitics and Bermúdez-Otero and Payne 2011 for one of special clitics). According to Kroeger's (2005:322f) systematization, simple clitics are further subdivided in bound words like the reduced auxiliaries of English (e.g. *I've, he's*) and Indonesian = *nya* '3SG' (cf. Cohn 1989) on the one hand and phrasal affixes like English "genitive *'s*" (e.g. *the Queen of England's cat*) on the other. Special clitics also come in two well-known guises: verbal clitics (e.g. Romance person markers like *la* 'her' in Spanish *la quiero ver* = *quiero verla* 'I want to see her') and second-position clitics (e.g. the Russian polar question marker *li* and Czech person markers like *ji* 'her' in *vidím ji* 'I see her' and *dnes ji vidím* 'I see her today'). Further special clitics include floating clitics (like Latin interrogative *ne* in *hicne vir usquam nisi in patria morietur?* 'shall this man die anywhere but in his native land?') and clause-final clitics (like the Iquito enclitic *na*, which marks finite subordinate clauses, cf. Beier 2004).

In a recent study, Spencer and Luís (2010) see words and affixes as polar opposites and work with a canonical definition, according to which clitics are formally like a canonical affix (i.e., they are prosodically unspecified and dependent on another element and consist of a monomoraic CV syllable) but distributionally like a canonical function word (i.e., they occur at the edge of the phrase—rather than a word—that is its morphosyntactic domain and have wide scope over coordinated phrases). This naturally allows for a number of non-canonical elements that show properties of clitics and affixes in varying degrees. Second-position clitics, e.g., are like canonical clitics with respect to form but show a distribution pattern of their own. The same applies, according to Spencer and Luís, to French definite articles, Turkish verb inflections, and Polish floating/mobile inflections (e.g. = *m* '1SG' and = *śmy* '1PL'), all of which are formally like canonical clitics but distributionally different from them, and different from each other.

Summing up, some of the parameters along which subtypes of clitics differ are syntagmatic while others are paradigmatic. The former include host selectivity, host type, position in the word/phrase/clause, and phonological cohesion values; the latter include whether they belong to a particular part of speech and whether they alternate with other elements like suffixes or words.[3] In a spirit similar to Spencer and

---

[3]Aikhenvald (2002) lists these and other formal properties, as well as some functional properties, including segmental properties, degree of cohesion, combinability with other clitics, scope, and lexicalization.

Luís's, Van Gijn and Zúñiga (2014) propose four parameters for the classification of morphological units: (i) phonological integration (yes/partial/no), (ii) positional rigidity (yes/no), (iii) host sensitivity to syntactic weight (yes/no), and (iv) host sensitivity to lexical category (yes/no). I will try to ascertain such properties in order to characterize different elements in Mapudungun that are candidates for clitichood in Sects. 3 through 5. Note in passing that the criticisms mentioned in Haspelmath (2011) with respect to the notion of grammatical or (morpho-)syntactic word as a crosslinguistically valid notion do not apply here because the analysis is explicitly language-specific. Mapudungun-specific p-words and g-words will be assumed to exist as separate from smaller (feet and morphemes) and larger units (phrases) in what follows; a justification of such an assumption lies beyond the scope of the present paper.

3 The question particle *am*

Mapudungun particles have not been the object of much attention in the older literature on the language, and Smeets (2008:Chap. 32) is the only study I am aware of that addresses them systematically, albeit sketchily. Not only are the contexts of occurrence and exact functions of these elements still poorly understood, but very little is also found on their formal characteristics. The present section will outline the basics of the particle *am* with respect to form and function in interrogative sentences.[4]

The examples in (1) illustrate its use with question words (a–c) and in polar questions (d).[5] Smeets says that *am* is employed here "to ask for more precise information, to express surprise or to provoke" (Smeets 2008:330). As usual in comparable cases in other languages, *am* is not strictly obligatory but often regarded by speakers as making (some) questions more idiomatic. Approximate translational equivalents seem to be German *denn* or *etwa* and Spanish *acaso*; in English, *really* seems to cover part of its functions.

(1)   *am* 'Q' (Smeets 2008:330–331, Augusta 1916:76)

   a.   *Chumngechi am dewmayafiñ?*
        chumngechi=am dewma-a-fi-n
        how=Q   make-FUT-3.OBJ-1SG.SBJ.IND
        'How shall I make/finish it?' (a question to the person in charge)
        (cf. *chumngechi Ø dewmayafiñ*—the speaker wonders to him/herself)

---

[4]This particle (or its allomorph *kam*) also occurs in affirmative sentences, sometimes with a clear causal function (e.g. *amulayan iñche, füchalu kam* [I.will.not.go 1SG get.old.PCPL PART] 'I will not go, because I am [too] old') and on occasion with a focus-like yield (*eymi kam mütẹn* [2SG PART only] 'it is only up to you (SG)') (Augusta 1916:76). A similar semantic connection is found in Germanic: High German *denn* and *dann* alternated more or less freely until the 18th century (now, *denn* means 'because, than' and also occurs as an interrogative particle, and *dann* means 'then'). The German elements are cognates of English *then* and *than*, which also alternated freely until ca. 1700.

[5]The same *am* seems to occur in compound particles as well (as enclitic in *chiam* < *chi = am* and as proclitic in *anta* < *am = ta* and *anchi* < *am = chi*), but such elements, as well as their building blocks *chi* and especially *ta* ~ *ṭa*, are even more poorly understood (cf. Smeets 2008:335f).

    b.   *Chem mew am femllenuafulu?*
         chem mew=am fem-lle-nu-a-fu-lu
         what POSTP=Q do.so-CERT-NEG-FUT-RI-PTCP
         'Why shouldn't [they] have done that?' ("of course they did!")

    c.   *Fey chumal am rekülamapafimi?*
         fey   chum-a-el=am   rekülama-pa-fi-i-m-i
         then be.what-FUT-NFIN=Q claim-CIS-3.OBJ-IND-2-SG
         'Then what have you (SG) come to claim it for?'

    d.   *Nepeley ñi püñeñ am?*
         nepe-le-i-Ø   ñi   püñeñ=am?
         wake.up-RES-IND-3 1SG.PSR child.of.woman=Q
         'Is my child awake?' (asks the mother, hearing sounds in the house)

Am attaches at the end of a constituent and actually attracts primary stress—simply following the stress allocation rules of the language: it is (part of) a final closed syllable. The reasons to say that it attaches to a host are two: (i) it never occurs as an isolated utterance, or after a pause, and (ii) together with the preceding element, it forms a new p-word for purposes of stress (cf. *chumal* [ʧuˈmal] vs. *chumal am* [ʧumaˈlam] in (1c)). There seems to be a clear tendency for *am* to occur immediately after the question word (which can be plurimorphemic, as *chem mew* 'why' (lit. 'for what') in (1b)) and clause-finally in polar questions. Although the issue merits exploration in more detail, my data suggest that complex interrogative sentences strongly disfavor *am*-marking. In diminishing order of preference, the admissible possibilities are only simple main clause and subordinate clause followed by *am*-marked main clause. Speakers are clearly more than uncomfortable with attempts to have =*am* appear at the end of a subordinate clause following a main clause, and I have not found such cases in published texts.

    The element *am* has no free-standing counterpart, nor a clear part of speech it might be assigned to. It might be regarded as simultaneously host-neutral and positionally free (i.e., as a phrasal affix), but this would neglect the fact that its occurrence with polar and nonpolar questions seems to be in systematic complementary distribution. With respect to Spencer and Luís's canonical approach, *am* is prosodically dependent but does not conform to the CV canon and is therefore not entirely formally canonical; nevertheless, it is distributionally canonical (at the edge of its relevant phrase, which is different for polar and nonpolar questions, and showing wide scope).

4 Adpositions

The prepositions most commonly used in the language are the ones listed in (2):[6]

---

[6]I am glossing over several details here, e.g. there is a homonymous adnominal plural marker *pu*, *ina* also occurs as a verb root meaning 'follow, chase', and *wente* and *miñche* are originally adverbs and used to cooccur with *mew* (see further down in the main body of text). See Augusta (1903:128f) for other prepositional elements.

(2)   Prepositions
  a.   *pu* 'in'   e.g. *pu ruka*   (in house)   'in the house'
  b.   *ina* 'next to'   e.g. *ina rüpü*   (next.to road)   'next to the road'
  c.   *wente* 'on'   e.g. *wente mesa*   (on table)   'on the table'
  d.   *miñche* 'under'   e.g. *miñche ruka* (under house)   'under the house'

Prepositions form a closed lexical class and are positionally restricted: they always appear before an NP, which minimally consists of a simple noun root or a personal pronoun but can be as complex as a multiradical nominal with derivational suffixes and preceded by adjectives, demonstratives, articles, and an embedded possessor NP. Prosodically, they can be the most prominent element of the PP only in the context of contrastive focus. They do not appear in a reduced form but have both a proclitic and an autonomous form.

The postpositions most commonly found are given in (3), viz. *püle* 'toward' (a), *kütu* 'from' (b), and the semantically unspecified *mew* (c):[7]

(3)   Postpositions

  a.   *l̲afke̲n mapu püle*
       sea   land   toward
       'toward the coast'

  b.   *tüfa kütu Troltren fentren   mapu*
       this from T.   as.far.as land
       'from here to Toltén'

  c.   – *Troltren* (*mew*) *amuan*
          T.   POSTP I.will.go
          'I will go to Toltén'
       – *Troltren* (*mew*) *küpan*
          T.   POSTP I.came
          'I came from Toltén'
       – *Anton ñi   furi   mew*
          A.   3.PSR back POSTP
          'behind Antonio'
       – *Aneltufiñ kiñe kuchillo mew*.
          aneltu-fi-n   *kiñe kuchillo mew*
          threaten-3.OBJ-1SG.SBJ.IND one knife   POSTP
          'I threatened him with a knife.'

Like prepositions, postpositions form a closed lexical class and are positionally restricted: they always appear after an NP. Prosodically, they can be the most prominent element of the PP only in the context of contrastive focus—which seems to be significantly less common than with prepositions. They appear in either a full or a reduced form: *püle* 'toward' as either [pɨˈle] or [pəˈle], *kütu* 'from' as either [kɨˈtu] or [kəˈtu] ∼ [ktu], and *mew* as either [mew] or [mo] ∼ [mu].[8] Note that, unlike the former two adpositions, *mew* does not have a disyllabic variant. This is relevant for

---

[7]*Mew* is often omitted as syntactic licenser of location/source/goal NPs with motion or posture verbs.

[8]Smeets (2008:61f) labels *mew* ∼ *mu* as "instrumental case suffix."

the stress placement rules, as shown in the variation illustrated in (4); since the unit consists of one syllable, it can attract stress when closed (*mew*) but usually does not when open (*mo* ∼ *mu*):

(4)   The (morpho-)phonology of *mew*

    a.   *chi rüpü mew* 'on the road'
        [ʧiɨˈpəmu] ∼ [ʧiɨˈɭpəmew] ∼ [ʧiˈɭəpɨˌmew]

    b.   *chi kuchillo mew* 'with the knife'
        [ʧikuˈʧiʎumu] ∼ [ʧikuˈʧiʎomu] ∼ [ʧikuˈʧiʎumew] ∼ [ʧikuˈʧiʎuˌmew]
        ∼ [ʧikuˈʧiʎoˌmew]

    c.   *chi wentru mew* 'for/to the man'
        [ʧiˈwentṣumu] ∼ [ʧiˈwentṣumew] ∼ [ʧiˈwentṣuˌmew]

This postposition can be realized either as the final syllable of a new p-word or as a separate, non-prominent (i.e. secondarily stressed) p-word in its own right. In principle, the [o] ∼ [u] alternation can be seen as indicating a word boundary: word-final /o/ can be realized as [u] in allegro speech, as in *wentro* or *wentru* [ˈwentṣu] 'man', but word-initial /o/ never is. (There are exceptions, like *ko* 'water', which is never reduced to **ku*.) There is considerable variation, however; my consultants realized *kuchillo* 'knife' variably before *mew* (4c), but tended to consistently realize *wentru* with a final *u*.

    I close this brief outline by concluding that Mapudungun adpositions in general, and the postposition *mew* in particular, have both a clitic and an autonomous form. Phonologically, proclitics seem to be a bit more autonomous (i.e. less reduced) than enclitics. As to their (non)canonical properties, the bound variants are positionally rigid and do appear at the edge of the relevant morphosyntactic domain (which can be either light or heavy NPs), but they do not show clear wide scope over coordinated NPs (e.g. ?*kuchillo ka toki mew* seems to be possible but dispreferred when compared to *kuchillo mew ka toki mew* 'with a knife and an axe').

5 Selected 3rd person markers of Mapudungun

This section covers two distinct phenomena. The first one is a set of markers of non-singular 3rd person participants that occur with varying degrees of attachment to other elements, and the second is a marker *mew*—apparently homophonous with the postposition presented in the preceding section—that encodes 3rd person actors in inverse finite verb forms. The interest of such phenomena lies in the fact that they all index verbal arguments or possessors but show disparate formal and distributional properties that have not been coherently systematized in the literature. I will address them in turn.

5.1 Nonsingular third person markers

The three persons and three numbers distinguished in Mapudungun are most clearly seen on the personal and possessive pronouns listed in Table 1 below. The basic analysis of these forms is that there are four specialized stems (*iñch-* '1', *ei-* '2', distal

Table 1 Personal and possessive pronouns in Mapudungun

	Personal (preverbal)	Personal (postverbal)	Possessive
1SG	iñch-e > iñche		ñi ___
1DU	iñch-i-u > iñchiw		i-u > yu ___
1PL	iñch-i-n > iñchiñ		i-n > iñ ___
2SG	ei-m-i > eymi		m-i > mi ___
2DU	ei-m-u > eymu		m-u > mu ___
2PL	ei-m-n > eymün		m-n > mün ___
3SG/3INAN	fey		ñi ___
3DU.ANIM	fey+eng-u > feyengu	+eng-u > (y)engu	ñi ___ engu
3PL.ANIM	fey+eng-n > feyengün	+eng-n > (y)engün	ñi ___ engün

demonstrative *fey*, and *e-* '3NSG')[9] and several inflectional suffixes (-*i* '1', -*m* '2', -*Ø* '3'; -*i* 'SG', -*u* 'DU', -*n* 'PL'; the portmanteau forms *ñi* and -*e* for the 1st (and 3rd) persons are non-compositional). While possessive markers always precede their possessed nominal (cf. *ñi ruka* 'my/his/her house', or *Juan ñi ruka* 'Juan's house'), personal pronouns can either precede or follow the verb form to which they correspond; 3rd person nonsingular forms have preverbal and postverbal allomorphs. Interestingly enough, and contravening the default stress assignment rules in the language, the final *u* on the 3rd person dual forms attracts stress. The symbol "+" in what follows represents a morpheme boundary unspecified for affixhood/clitichood status; the nature of such boundary will be the topic of discussion further down.

Mapudungun verb forms can be finite, marking at most two (bipersonal) and at least one participant (monopersonal) or nonfinite, which mark one participant less than their finite counterpart and have some restrictions as to their TAM inflectional potential. The suffixes marking person and number of the privileged syntactic argument or "subject" of finite forms are virtually identical to those marking those categories on pronouns, cf. Table 2 below; additionally, these forms take a mood marker (-*i* 'IND', -*l* 'SUBJ', -*Ø* 'IMPER'). In such paradigms, the final *u* on the 3rd person dual forms attracts stress as well:

The examples in (5) below illustrate the typical use of pronouns and indicative verb forms. In (a), the 3rd person possessive pronoun *ñi* is used with a nonsingular possessor whose referent is retrievable from the context; *ñi chaw engu* or *ñi chaw engün* would be used to explicitly express dual or plural possessors respectively. Personal pronouns are usually not necessary in main clauses and often only verb morphology expresses person (b). Nevertheless, personal pronouns can be added for focus/disambiguation purposes. In (c), the 3rd person dual object is indexed by -*fi* on the (direct bipersonal) verb, which is unspecified for number, so the element *engu* can be encliticized to disambiguate. Similarly, in (d), the 3rd person dual subject is

---

[9]Smeets (2008:99) hypothesizes *ə-n-che > iñche*, *ə-n-che-u > iñchiw*, and *ə-n-che-iñ > iñchiñ*, where *n* is the 1SG.IND marker (*n-ch > ñch*), *ə* is fronted (and raised) before *ñ*, and * *che* 'person' is equivalent to the current lexical item. In her view, the element *e* in 2nd person pronouns "is [also] probably related" to inverse -*e*. I remain unconvinced (but agnostic) with respect to these proposed etymologies.

Table 2 Monopersonal verb forms in Mapudungun (underlying, simplified; *tripa-* 'exit')

	Indicative	Subjunctive	Imperative
1SG	tripa-n > tripan	tripa-li > tripali	tripa-chi > tripachi
1DU	tripa-i-i-u > tripaiyu	tripa-l-i-u > tripaliyu	tripa-Ø-i-u > tripayu
1PL	tripa-i-i-n > tripaiyiñ	tripa-l-i-n >tripaliñ	tripa- Ø-i-n > tripaiñ
2SG	tripa-i-m-i > tripaymi	tripa-l-m-i > tripalmi	tripa-nge > tripange
2DU	tripa-i-m-u > tripaymu	tripa-l-m-u > tripalmu	tripa-Ø-m-u > tripamu
2PL	tripa-i-m-n > tripaymün	tripa-l-m-n > tripalmün	tripa-Ø-m-n > tripamün
3SG	tripa-i-Ø	tripa-le	tripa-pe
	>tripay[a]	>tripale	>tripape
3DU	tripa-i+ng-u	tripa-l+ng-u	tripa-pe+eng-u
	>tripaingu	>tripalngu	>tripape yengu
3PL	tripa-i+ng-n	tripa-l+ng-n	tripa-pe+eng-n
	>tripaingün	>tripalngün	>tripape yengün

[a]The syllabic vs. nonsyllabic status of the indicative marker in 3rd person forms is less straightforward than the spelling conventions *-i* and *-y* may suggest. (And there is considerable variation as to how glides are orthographically rendered in Mapudungun studies.) Postvocalically and verb-finally, the indicative morpheme is always realized nonsyllabically (e.g. *tripa-i > tripay* 's/he exited'; after *i* it is simply deleted, e.g. *pi-i > pi* 's/he said'). If the preceding segment is a consonant or a glide, it is always the nucleus of the word-final syllable (e.g. *kon-i > koni* 's/he entered', *küdaw-i > küdawi* 's/he worked'). Salas (2006:88f) explicitly says that "*y* turns to *i* before *engu* or *engün*" in *tripay*, but the same source notes the fact that an alternative pronunciation is possible as well: even though four-syllable [ʈʂi.ˌpa.i.ˈŋu] is probably the default realization, three-syllable [ʈʂi.ˌpai̯.ˈŋu] is also available in allegro speech (the stress pattern indeed suggests that it is a reduced form). The fact that *i* does become nonsyllabic before either 2nd person *m* or 3rd person *mew* (cf. Sect. 5.2) suggests that the appended elements may have an underlying/etymological *e* and do not/did not start with *ng*

marked by a zero on the verb, which is unspecified for number in inverse bipersonal forms, so the corresponding enclitic *engu* is appended.[10] Interestingly, (e) shows that such enclitics are not strictly verbal; rather, they appear clause-finally. Lastly, a pre-verbal pronoun can be seen in (f).

(5) Indicative verb forms and pronouns/NPs (Salas 2006:*Latrapay*.2–3; Smeets 2008:98)

   a. *Mongelekefuy ñi chaw.*
       monge-le-ke-fu-i-Ø   ñi   chao
       live-PROG-HAB-RI-IND-3 3.PSR father
       'Their father was alive then.'

   b. *Kuñifallngeingu.*
       kuñifall-nge-i+ng-u
       abandon-PASS-IND+3NSG-DU
       'They were both left behind/abandoned.'

---

[10]See Zúñiga (2006:Chap. VII) for a detailed account of direct and inverse verb forms in Mapudungun.

    c.   *Lelifiñ engu.*
        leli-fi-n+eng-u
        look.at-3.OBJ-1SG.SBJ.IND+3NSG-DU
        'I watched the two of them.'

    d.   *Fey ṭa ngünayekefeyew engu.*
        fey   ṭa
        DEM[SG] PART
        ngüna-ye-ke-fu-e-i-Ø-mew+eng-u
        protect-DISTR-HAB-RI-INV-IND-3.SBJ-3.OBJ+3NSG-3DU
        'This [one] used to protect them both.'

    e.   *Aṉümkay pinu yengün.*
        aṉü-m-ka-i-Ø   *pinu+eng-n*
        sit.down-CAUS-CONT-IND-3   cane+3NSG-PL
        'They (PL) planted cane.'

    f.   *Feyengu akuingu.*
        Fey+eng-u   *aku-i+ng-u*
        DEM+3NSG-DU arrive.here-IND+3NSG-DU
        'They (DU) arrived here.'

    g.   *Akuy (*ngu) epu wentru.*
        aku-i(-Ø)+ng-u   *epu wentro*
        arrive.here-IND-3+3NSG-DU two man
        'Two men arrived here.'

Observe in (5g) above that lexical nonsingular 3rd persons are in complementary distribution with the verbal marking: both *akuy epu wentru* 'two men arrived here' and *akuingu* 'they both arrived here' are possible, but **akuingu epu wentru* is not. With pronominal 3rd persons, there is some variation; *feyengu akuingu* seems to be clearly preferred to *feyengu akuy* (5f above) (which is also possible); in the plural both *feyengün akuingün* and *feyengün akuy* seem to be roughly equally acceptable. Note in this context the following examples from Salas (2006:222–223):

(6)   (*Piwi ṭa epu kuñifall wentru. Amutuingu.*)
     ('[Thus] the two poor men said to each other. Both of them went away.)
     Amuingu ka powi kiñe fütra witrulku mew engu.

     amu-i+ng-u   *ka* *pow-i-Ø*   *kiñe fütra*
     go-IND+3NSG-DU   and arrive.there-IND-3 one big
     witrulku mew+eng-u
     brook   POSTP+3NSG-DU
     'They (DU) set off and arrived at a big brook.'

The lexical NP *epu kuñifall wentru* 'two poor men' establishes reference, number, and other features of the main participants in a preceding clause; the next clause (*amutuingu* 'they (DU) went away' shows the expected number marker on the verb). The first verb form of the last clause (*amuingu* 'they (DU) went') carries the 3rd person dual marker, whereas the second verb form (*powi* 'X arrived there') is unmarked for number; the 3DU marker *engu* comes at the very end of the clause, in its postverbal form as expected.

A further use of the personal pronoun is to mark accompaniment, as shown in (7).[11] Unsurprisingly, the form *engu* is used for two participants (a) and *engün* is used with more than two (b):

(7) "Comitative use" of the personal pronoun (Smeets 2008:139–140; Salas 2006:90)

  a. *Alü kamapuley Peru engu Argentina.*
      alü   kamapu-le-i-Ø   Peru eng-u   Argentina
      much far.away-RES-IND-3 P.   3NSG-DU A.
      'Peru and Argentina are far away.'

  b. *Kom ñi pu che engün kañpüle amuy.*
      kom ñi   pu che   eng-n   kañpüle   amu-i-Ø
      all   3.PSR PL person 3NSG-PL elsewhere go-IND-3
      'He went elsewhere with his entire family.'

  c. *Antonio amuay Temuko ñi pu karukatu engün.*
      Antonio amu-a-i-Ø   Temuko ñi   pu karukato eng-n
      A.   go-FUT-IND-3 T.   3.PSR PL neighbor 3NSG-PL
      'Antonio will go to Temuco with his neighbors (PL).'

The synchronic description of these nonsingular 3rd person markers has to distinguish several kinds of markers due to their formal make-up and distribution. I use Van Gijn and Zúñiga's (2014) systematization in Table 3 (also referring to Spencer and Luís's 2010 scope parameter in the text below).

The *-ngu ∼ -ngün* markers on indicative and subjunctive verb forms qualify as canonical affixes and are actually treated as affixes in the literature on Mapudungun; they show maximal cohesion with their host, which cannot but be a finite verb form, and they do not show wide scope even although they stand in something of a complementary distribution with lexical NPs. Most other elements resemble function words regarding their form: they are more autonomous and not reduced. Nevertheless, not all of them conform to the canonical function-word pattern regarding their distribution. The nonadjacent postverbal personal pronoun would come closest to wordhood (evidence is inconclusive at this point; some speakers prefer one such element per verb form, with some tendency towards appending it to the predicate, while others—perhaps most—treat such elements as more loosely attached); Mapudungun studies have commonly treated them as "pronouns," i.e., as separate words. The possessor explicator behaves mostly alike (*?ñi ñuke ka ñi chaw engu* [3.PSR mother and 3.PSR father 3DU] 'their (DU) mother and father' is possible in principle but strongly dispreferred; alternatives like *ñi ñuke engu ka ñi chaw engu* or *feyengu ñi ñuke ka ñi chaw* are favored by speakers), but it must follow an NP. The coordinating postposition attaches to an NP but does not show wide scope, i.e., it has to be repeated after each coordinated element. Most interestingly, the imperative 3NSG markers and the element appended to the demonstrative are affix-like with respect to their distribution (lexical host selectivity and narrow scope), but they seem to retain considerable formal autonomy. Observe that only the imperative 3NSG markers can optionally take

---

[11] Augusta (1903:128f) treated such elements as equivalent to Spanish *con* 'with', i.e., as "prepositions."

Table 3 Selected properties of nonsingular 3rd person markers (e.g. … *ng-u* '3DU')

	Phonological integration	Positional rigidity	Syntactic weight	Lexical class
Nonimperative finite marker: *amutuingu* 'they (DU) went away'	✓	✓	✓	✓
Nonadjacent postverbal marker: *amutuy … engu* 'they (DU) went away'	✗	✓	✗	✗
Possessor explicator: *ñi chaw engu* 'their (DU) father'	✗	✓	✗	✓
Coordinating postposition: *Chile engu Peru* 'Chile and Peru'	✗	✓	✓	✗
Imperative finite marker: *tripape (y)engu* 'may they (DU) exit'	(✓)	✓	✓	✓
Preverbal pronoun: *feyengu amutuy* 'they (DU) went away'	(✓)	✓	✓	✓

a connecting glide *y* if the preceding segment is a vowel. Furthermore, stress patterns differ, suggesting that the imperative 3NSG markers, but not adnominal *engu* ∼ *engün*, are part of one p-word; the latter are stressed themselves and do not change the stress of the preceding (or following) word (e.g., *rúka* 'house', *máñke* 'condor', *pewén* 'monkey puzzle tree', and *tralkán* 'thunder' will normally retain their stress before/after *engu* or *engün*). The imperative 3NSG markers, by contrast, do change the stress pattern (e.g., *pínu* [ˈpinu] 'cane', with a stressed penult, normally becomes *pinù yengú* [piˌnuyeˈŋu], with secondary stress on the antepenult and primary stress on the ultima in (5e) above).

This leaves us with an interestingly complex picture. Nonsingular 3rd person markers are located along a multidimensional continuum of phonological and distributional properties: between the two poles of affixal and (almost fully) word-like markers, we find possibly two distinct non-canonical types of non-autonomous elements. Both resemble function words regarding form but are either like French definite articles or like affixes regarding distribution.

5.2 Third person actor marker with inverse finite verb forms

Mapudungun transitive verbs can appear in direct and inverse forms. Roughly, the former are used when 1st or 2nd persons act on 3rd persons, or when a more prominent 3rd person acts on a less prominent one; the latter are employed when a 3rd person acts on a speech act participant, or when a less prominent 3rd person acts on a more prominent one. Finite inverse forms, therefore, mark two persons: the prominent (patientive) one, distinguishing both person and number (the "subject"), and the

less prominent (agentive) one, which is simply 3rd person (the "object"). In addition, these inverse forms take the suffix *-e*; an indicative paradigm with *pi-* 'say, tell' can be seen in (8) below. Note that those with singular subjects lose some segments in their surface forms; *mew* turns to *ew* with 1SG and 3SG, and the sequence *mimew* is reduced to *mew* with 2SG:

(8) Verbal *mew* (with inverse indicative forms of *pi-* 'say, tell')

3→1SG	*pi-e-n-mew* >	*pienew*
3→1DU	*pi-e-i-i-u-mew* >	*pieyumew*
3→1PL	*pi-e-i-i-n-mew* >	*pieiñmew*
3→2SG	*pi-e-i-m-i-mew* >	*pieymew*
3→2DU	*pi-e-i-m-u-mew* >	*pieymumew*
3→2PL	*pi-e-i-m-n-mew* >	*pieymünmew*
3→3SG	*pi-e-i-Ø-mew* >	*pieyew*

Similar reductions occur with most singular subjects in the subjunctive and with all (unmarked) subjects in some nonfinite forms.[12]

Thus, the element encoding 3rd person actors-objects on inverse verbs is (*m*)*ew*.[13] Even though inverse *-e* attracts stress in some, but not all, of these forms, this (*m*)*ew* is best analyzed as a verbal suffix. Even if it were a cognate of the postposition presented in Sect. 4, it differs from it rather markedly both in formal and functional terms. In Spencer and Luís's terms, this *mo* ∼ *mu* qualifies as canonical affix; its non-strictly-CV allomorph *mew* is almost canonical.

6 Summary and conclusions

The present paper outlined some parameters of the formal make-up and behavior of three functionally different dependent elements, viz. the interrogative marker *am*, a number of adpositions, and several 3rd person indexes. In terms of the typology sketched in Sect. 2, *am* is a simple clitic of the phrasal affix subtype, a next-to-canonical clitic; it is probably best to assign particle status to it, and it has no counterpart that differs from it in either form or distribution. The semantically unspecified postposition *mew*, by contrast, is a function word that can occur (occasionally, or dialectally, reduced to *mu*) as a simple clitic of the bound word subtype, or as a non-canonical clitic. The indexes of 3rd person participants are more varied. Glossing over the details of how dependence is parameterized merely for the sake of expository clarity, the overall picture is schematically summarized in Table 4.

---

[12]Singular subject subjunctive inverses are 3→2SG *pi-e-l-m-i-mew* > *pielmew* (instead of *pielmimew), 3→3SG *pi-e-le-mew* > *pieliyew* (instead of *pielemew), and nonreduced 3→1SG *pi-e-li-mew* > *pielimew*. Nonfinite forms replace the mood-person-number sequence of finite forms with a nonfinite ending like *-n*, *-el*, *-am*, *-lu*, etc. It appears plausible to reconstruct the inverse nonfinite counterpart *-etew* of direct *-el* as **-e-et-mew*, consisting of the inverse marker **-e*, the nonfinite marker **-et* (which appears as *-el* word-finally, by a general rule that precludes syllable-final plosives from appearing in the language), and the 3rd person actor marker **-mew*.

[13]Eastern varieties of Mapudungun, which have *mu* as the default realization of the postposition, also have (*m*)*u* as the default realization of this marker.

Table 4 Selected Mapudungun elements and their properties

	Formal dependence	Formal independence
Distributional dependence	Affixes: – nonimperative *-ng-*{NUM} – verbal *-mew* ∼ *-mu*	"Anti-clitics": several 3NSG markers
Distributional independence	Clitics: – interrogative =*am* – postposition =*mew* ∼ =*mu*	Words: – postposition *mew* – nonadjacent postverbal *eng-*{NUM} – possessor explicator *eng-*{NUM}

The preceding section gave a more detailed account of the several "anti-clitics" identified for Mapudungun, i.e. units that are probably best seen as (largely) independent p-words (or, at least, as units that do not interfere with the host p-word in terms of stress) but as parts of larger g-words. Dixon and Aikhenvald (2002:28) mention comparable elements in Yidiny, Jarawara, Georgian, and Turkish. To my knowledge, such elements have not been studied in great detail from a crosslinguistic perspective. Dixon's examples, Spencer and Luís's comments on French, Turkish, and Polish non-canonical clitics mentioned in Sect. 2 above, and some of the Mapudungun 3NSG markers addressed in Sect. 5, however, suggest that any theory of cliticization in particular and of wordhood in general would benefit from exploring different parameters of formal and distributional dependence in similar phenomena on a wide basis. Binary approximations like dependence vs. independence are best replaced by fine-grained parameter sets like the ones used in Zwicky (1985), Aikhenvald (2002), Spencer and Luís (2010), and Van Gijn and Zúñiga (2014), in order to arrive at a more comprehensive picture. It is clearly an empirical question whether some, many, or even most of the most powerful parameters lead to robust clusters (cf. Haspelmath's 2011 critique of the crosslinguistically valid notion of g-word).

As far as the description of Mapudungun is concerned, the language does not seem to have special clitics of either the Romance-like verbal clitic type or the second-position type. In fact, and even though clear cases of words and affixes exist, the elements discussed in this study suggest that several intermediate values of the dependence parameters are not only possible but perhaps also stable, with respect to both their (anti-)clitic quality and their non-canonicity (and non-prototypicality). Degrees of cohesion and kinds of cohesion are dimensions that need more in-depth explorations, both language-specifically and crosslinguistically, irrespective of the particular theory of wordhood that one would rather espouse.

Abbreviations and symbols ANIM animate, CAUS causative, CERT certainty, CIS cislocative, CONT continuative, DEM demonstrative, DISTR distributive, DU dual, FUT future, HAB habitual, IMPER imperative, INAN inanimate, IND indicative, INV inverse, NEG negation, NFIN nonfinite, NSG nonsingular, NUM number, OBJ object, PART particle, PASS passive, PCPL participle, PL plural, POSTP postposition, PROG progressive, PSR possessor, PTCP participle, Q question, RES resultative, RI ruptured implicature, SBJ subject, SG singular, SUBJ subjunctive

\-   affixation boundary
=   cliticization boundary
prosodic word boundary
+   unspecified morpheme boundary
{X}   set of different values/forms for *x*
x→y   '*x* acting on *y*'

References

Aikhenvald, A. (2002). Typological parameters for the study of clitics, with special reference to Tariana. In R. M. W. Dixon & A. Aikhenvald (Eds.), *Word. A cross-linguistic typology* (pp. 42–78). Cambridge: Cambridge University Press.

Anderson, S. (2005). *Aspects of the theory of clitics*. Oxford: Oxford University Press.

de Augusta, F. J. (1903). *Gramática mapuche bilingüe*. Santiago: Séneca. [1990].

de Augusta, F. J. (1916). *Diccionario araucano: mapuche-español, español-mapuche*. Santiago: Cerro Manquehue. [2007].

Beier, C. (2004). El marcador clítico de fin de cláusula -*na*. Ms., Iquito Language Documentation Project, University of Texas at Austin.

Bermúdez-Otero, R., & Payne, J. (2011). There are no special clitics. In A. Galani, G. Hicks, & G. Tsoulas (Eds.), *Morphology and its interfaces* (pp. 57–96). Amsterdam: Benjamins.

Cohn, A. (1989). Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory*, *7*(2), 167–216.

Dixon, R. M. W., & Aikhenvald, A. (2002). Word: a typological framework. In R. M. W. Dixon & A. Aikhenvald (Eds.), *Word. A cross-linguistic typology* (pp. 1–41). Cambridge: Cambridge University Press.

Echeverría, M., & Contreras, H. (1965). Araucanian phonemics. *International Journal of American Linguistics*, *31*, 132–135.

Van Gijn, R., & Zúñiga, F. (2014). Word and the Americanist perspective. *Morphology*, September 2014.

Goedemans, R., & van der Hulst, H. (2013). Fixed stress locations. In M. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online*, Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/14, Accessed on 2013-12-12).

Halpern, A. (1998). Clitics. In: Spencer & Zwicky (Eds.) (pp. 101–122).

Haspelmath, M. (2011). The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica*, *45*(1), 31–80.

Kroeger, P. (2005). *Analyzing grammar. An introduction*. Cambridge: Cambridge University Press.

Salas, A. (1978). Mapuche-español. Análisis fonológico contrastivo. *VICUS Cuadernos, Lingüística*, *II*, 57–86.

Salas, A. (2006). *El mapuche o araucano* (2nd, revised ed.). Santiago: Centro de Estudios Públicos.

Smeets, I. (2008). *A grammar of Mapuche*. Berlin: de Gruyter.

Spencer, A., & Luís, A. (2010). The canonical clitic. In D. Brown, M. Chumakina, & G. Corbett (Eds.), *Canonical morphology and syntax* (pp. 123–149). Oxford: Clarendon Press.

Spencer, A. & Zwicky, A. (Eds.) (1998). *The handbook of morphology*. Oxford: Blackwell.

Zúñiga, F. (2006). *Mapudungun—el habla mapuche*. Santiago: Centro de Estudios Públicos.

Zwicky, A. (1977). *On clitics*. Bloomington: Indiana University Linguistic Club.

Zwicky, A. (1985). Clitics and particles. *Language*, *61*(2), 283–305.

Zwicky, A., & Pullum, G. (1983). Cliticization vs. inflection: English *n't*. *Language*, *59*(3), 502–513.