On an Oddity in the Development of Weak Pronouns in Deictic Expressions in the Languages of the Balkans*

Brian D. Joseph
The Ohio State University

Preamble and Dedication

As is well-known, a number of the languages of Southeastern Europe — in particular Albanian, Bulgarian, Greek, Macedonian, Romanian, Romani, Serbian, and, to a lesser extent, Turkish — show numerous curiosities in their historical development that have long intrigued scholars and have led to their characterization as forming a "Sprachbund," an area of linguistic convergence in which several languages show structural similarities that are due to contact between and among their speakers and not inheritance from a common ancestral language. This paper is devoted to the explication of a particular Balkan linguistic curiosity regarding the occurrence of certain types of weak pronouns in deictic expressions that has not previously been noticed, and thus not previously explained. It is entirely fitting therefore that such a paper should appear in a volume devoted to honoring Howard Aronson, inasmuch as a considerable portion of his scholarly energy throughout his career, as well as the impressive output that resulted from these efforts, has been aimed at clarifying grammatical oddities, not just in the Balkans — such as the nature of the verbal system of Bulgarian — but in another hotbed of linguistic curiosity as well, namely the Caucasus region and especially the Georgian language found there. The remarks contained herein are thus dedicated to Howie in appreciation of the interesting work with which he has inspired us over the years.

1. Linguistic Preliminaries to the Problem

It is safe to say that virtually all languages, and perhaps indeed all,[1] show elements which are sometimes, for want of a better term, referred to as "little words,"

generally called "particles" and/or "clitics" in the literature,[2] which are typically short and prosodically deficient but at the same time are indispensable, since they often mark important grammatical categories or serve vital discourse functions, and yet are notoriously hard to classify. "Little words" play an important role in the striking grammatical parallels that hold among the Balkan languages, as for instance with the future marker that is a reduced form of (probably the 3SG present form) of 'want,' *cf.* (1a), with the postposed definite article, *cf.* (1b), with the element that combines with adjectives to form analytic comparatives, *cf.* (1c), or with the subordinating and modality marking element that figures prominently in Balkan complementation, what Friedman 1985 has called the Balkan "dental modal subordinator," *cf.* (1d):

(1) "Little Words" in Balkan Sprachbund Convergences

a. the future marker that is a reduced form of (probably 3SG of) 'want' (Albanian *do*, Bulgarian *£te*, Greek *θa*, Macedonian *k′e*, Romanian *o*, Romani *ka*, etc.)

b. the postposed definite article (*e.g.,* masc. nom. sg. Albanian *-i*, Bulgarian *-a`t*, Macedonian *-ot*, Romanian *-ul*, etc.)

c. the element that combines with adjectives to form analytic comparatives (*e.g.,* Albanian *më*, Bulgarian *po*, Greek *pjo*, Macedonian *po*, Romanian *mai*, etc.)

d. the "dental modal subordinator" (*e.g.,* Albanian *të*, Bulgarian *da*, Greek *na*, Macedonian *da*, Romanian *sa`*, Romani *te*).

Indeed, the large number of convergences of this sort that revolve around such prosodically weak elements has prompted at least one Balkanist, Klagstadt 1963, to suggest that there was a prosodic basis for much of what is seen in the way of common structures in the Balkan languages (and see also the work of Ronelle Alexander, *e.g.,* Alexander 1993, 1999).

     To be included in such a grouping of potentially prosodically based Balkan convergences that focus on "little words" are some involving weak pronominal

forms,[3] such as argument doubling (also known as object reduplication), illustrated in (2a), and the "dative" of possession with pronominal forms, illustrated in (2b):

(2) Balkanisms Involving Weak Pronominal Forms:

a. argument doubling (= object reduplication), *e.g.*, Alb *e pashë Gjonin* = Grk ***ton*** *îða ton jáni* = Mac ***go*** *gledav Ivan* = Rom *l'am va`zut pe Ion* 'I saw John' (literally, 'him I-saw (the-)John')

b. dative of possession (*e.g.,* BG *brat **mi*** 'my brother,' *Grk o fílos **mu*** 'my friend,' etc.)

Falling in further in such a categorization of Balkanisms is the convergence noted by Schaller 1975, mentioned as well in Banfi 1985, concerning weak accusative pronouns following a governing deictic expression, illustrated in (3):

(3)   Weak Accusative Pronominals in Deictic Expressions:

   a.   na ton
          DEICTIC him/ACC.WEAK
          'Here he is!' (Greek)

   b.   eto go
          DEICTIC him/ACC.WEAK
          'Here he is!' (Bulgarian)

In addition to the pattern of (3a) with an accusative, Greek has developed deictic expressions with nominative weak pronouns. These forms are given in (4), and their use is illustrated in (5); what is significant about them is that they occur with just the predicates **ná** 'here is/are' and **pún** 'where is/are?' as in (5a/b) but not with other verbs;[4] thus, (5c) and any conceivable sentence with these forms other than with the predicates **ná** and **pún** is effectively ungrammatical[5]:

(4)   Weak Nominative Forms in Greek

	MASC	FEM	NTR
SG	tos	ti	to
PL	ti	tes	ta

(5)   Use of Greek Weak Nominatives

    a.   ná tos
        DEICTIC he/NOM.WEAK
        'Here he is!'

    b.   p ún dos
        where-is he/NOM.WEAK
        'Where is he?'

    c.   *méni tos eðó
        lives/3SG he/NOM.WEAK here
        'He lives here.'
        *íne tos o filós mu
        is/3SG he/NOM.WEAK the-friend/NOM my
        'He is my friend.'

Interestingly, such forms have been recreated several times within the recent history of Greek, as shown by parallels in various Greek dialects, *e.g.,* that of Lefkas in the Ionian islands (Pernot 1934), suggesting that it is a fairly natural path of development. Moreover, Greek is not alone among Indo-European languages in this regard, for Hittite underwent a similar development in its prehistory.

Nonetheless, there is a curious fact about the Greek nominative weak pronouns with regard to the other Balkan languages; in particular, within the Balkans, Greek is unique in having innovated in this way,[6] and this is so despite the accusative parallels in deictics in other Balkan languages, and despite the recurrence of such nominatives in Greek dialects.

OF ALL THE SLAVS MY FAVORITES

This paper is aimed at explaining this oddity, and thus the relevant facts are developed in some detail in the sections that follow, and an answer is provided for why the creation of weak nominative forms in the Balkans is restricted to Greek and Greek alone.

2. Origin of tos *within Greek*

The emergence of **tos** must be understood in the context of the single most striking fact about it, namely that it is restricted in its occurrence to combining with **ná** and **p¿ún**, as noted above in (5). It is reasonable to posit, therefore, that the origin of **tos** is tied in some way to one of those predicates. It has been argued elsewhere (Joseph 1981, 1994, 2001, forthcoming) that looking to **ná** provides the best account of how **tos** developed, and in what follows, this scenario is summarized as a basis for understanding how **tos** remains unique among the weak pronominal systems of the Balkan languages.

It can be safely assumed that the original syntactic pattern found with **ná** was [**ná** + accusative]. Such a syntagm would occur in Greek if **ná** were a borrowing from Slavic (as suggested in Joseph 1981), since the combination of a deictic element with an accusative is widespread in South Slavic (Schaller 1975),[7] as illustrated in (3b) with Bulgarian. Alternatively, if **ná** is from an earlier Greek **e:ní**, abstracted out of **e:níde** (= **é:n** 'behold!' + **íde** (IMPV of 'see')), as suggested by Hatzidakis 1905 and endorsed by Andriotis 1983 (but questioned in Joseph 1981), an accusative complement would nonetheless be expected given the usual syntax of **íde**, which, as a verbal imperative, would govern an accusative object.

At that point, one has to assume that an innovative pattern of [**ná** + nominative] arose for full NP complements, probably from the reinterpretation of neuter nouns where the accusative and the nominative are identical — *cf., e.g.,* **ná to peðí** 'Here's the child' (NOM or ACC) — aided by semantics of deixis; note, in particular, that the post-**ná** noun functions at some level as a subject, for which nominative is the usual case. If it is assumed further that the third-person strong pronouns (MASC.NOM.SG **aftós**, MASC.ACC.SG **aftón**, etc.), which have the same distribution as ordinary nouns, could occur with **ná**, most likely under conditions of emphasis (*i.e.,* meaning 'Here he is!').

At this point, then, the [**ná** + nominative] and [**ná** + accusative] patterns would have co-existed for essentially the same meaning, a situation that continues

into contemporary Greek, where both nominative and accusative after **ná** are possible, as in (6):

(6)   Variability in **ná** Constructions in Modern Greek

      a.   **ná** ton yáni
           DEICTIC the-John/ACC
           'Here's John!'

      b.   **ná** o yánis
           DEICTIC the-John/NOM
           'Here's John!'

      c.   **ná** ton
           DEICTIC him/ACC.WEAK
           'Here he is!'

      d.   **ná** tos
           DEICTIC him/NOM.WEAK
           'Here he is!'

The presence of these variants and a recognition by speakers of a "mapping" between them could then have led to a proportional analogy that would have yielded the innovative form **tos**; this analogy is schematized in (7):

(7)   **ná** aftón   :   **ná** aftós   : :   **ná** ton   :   **ná** X, X —> tos
       ACC.STRONG   NOM.STRONG ACC.WEAK   NOM.WEAK

This analogy hinges on the possibility of a perceived relationship between the strong pronominal forms and the weak ones, such that WEAK = STRONG minus initial **af-** (*e.g.*, **aftón**/**ton**, in the accusative). Such a pattern is found as well with the genitive forms, *i.e.,* **aftú/tu** 'of him, his,' though admittedly these strong genitive forms are not common — more common are extended forms such as **aftunú**.

    While one might suppose that deriving **tos** just as a phonological reduction of the strong pronoun (*i.e.,* **ná aftós** —> **ná tos**), as suggested for the accusative

forms in standard accounts of Greek (*e.g.,* Browning 1983, Horrocks 1997), and as is held to be the case more generally in literature on "grammaticalization" regarding the origin of weak pronouns (*e.g.,* Haiman 1991), such an account runs into some problems. In particular, one might expect a result that is doubly accented (*i.e.,* ****ná tós**), and moreover, the reduction of an **-ft-** cluster to **-t-** is not a regular phonological process in Greek. It is true that Medieval Greek does show a 3SG strong pronominal form **atós** (beside **aftós**), which would take care of the problem with **-ft-**, and would give a basis for deriving **tos** directly via vowel contraction (**ná atós** —> **ná 'tos**), but such a source does not solve the double accent problem. Also, it does not generalize to dialectal forms such as Ionian **éntos** discussed below in *Section 3*, for ****énatos** would then be expected (there being no straightforward way of eliminating the **a-** of **atós** after **én**).[8] Moreover, this analogical account finds support in Balkan facts discussed below regarding the absence of **tos**-like forms elsewhere; that is, the most compelling solution to the Balkan curiosity concerning **tos** ties in well with the analogical origin of **tos**.

3. Parallel Developments Elsewhere

As noted above, the analogical development of **tos** with **ná** is not unparalleled, for similar developments leading to the creation of a weak nominative pronoun analogically have taken place elsewhere in other Greek dialects independently of what happened with **ná**, and outside of Greek, in Hittite as well. To the extent, then, that such developments are "natural" and not unexpected or surprising, the absence of parallels to **tos** within the Balkans, outside of Greek, becomes even more curious. By way of developing further this line of reasoning, the facts from other dialects of Greek and from Hittite are briefly reviewed here.

As indicated above, the dialect of Lefkas shows developments parallel to, yet apparently independent of, those giving **ná tos** elsewhere in Greek (Pernot 1934). Lefkas has a deictic element **é**, which is equivalent to Standard Greek **ná** 'here is,' and this occurs in, among other constructions, what Pernot calls a deictic "interjection" **é tos** 'Here he is,' with the same pattern as in the standard language. Unless this is simply based on the standard Greek **ná tos** construction with the substitution of the dialectally native deictic, presumably **é tos** is a result of the same sort of process as outlined above.

At first, though, there appears to be a small obstacle here, in that there is no *é ton, *i.e.,* [é + accusative], evident in Lefkas which would have been the basis for the creation of é tos. However, this absence need not be problematic, as *é ton could have been supplanted already by é tos by the early 20th century; still, given fluctuation still present in standard language between NOM and ACC with ná (see above, (6)), perhaps some trace of *é ton would be expected.

Interestingly, Lefkas shows a synonymous deictic "interjection" éntos (also found elsewhere in Ionian islands;[9] this too could show the results of the same process as led to ná tos, since an apparent accusative éntone occurs. This latter form is presumably *énton, with accusative, with extra -e so common in the latter stages of Greek as a "buffer" against weakening of final -n — *cf.,* the accusative weak pronoun variant tóne, for ton, even in the standard language — though to some extent this depends on the source of én.

If én is somehow related to ná — though it is not at all obvious what sort of processes might relate them — or instead derives directly from Ancient Greek én 'behold!' (though with shortening of é, not regular for Ionian), then *én with an accusative complement is as expected and thus positing *én ton, a basis for the independent creation of éntos, is not at all problematic. If, however, én is, as Pernot suggests, from éni, the Medieval Greek 3SG form of 'be,' then éntos would be primary, since nominative would be the expected case for the complement of 'be.' Presumably, this would be from én plus an already-created tos — perhaps borrowed from the standard language, for it is not clear how else it would have arisen[10] — and éntone would have been formed from it. Even though the creation of an apparent accusative éntone out of a nominative *éntos would involve the obscuring of the 'be' origin of the én- part, such a development is formally motivated analogically in a similar way to the creation of a nominative out of an accusative pattern, by just reversing the direction of the terms, so to speak, in the schema in (7). Moreover, the same type of counter-etymological reanalysis is not unparalleled, for it is what now allows for pún don in standard Greek for 'Where is he?,' presumably based on ná ton and pún dos (*cf.* (5b)), with an accusative that is anomalous, given the etymology of pún, from pú 'where?' plus a reduced form of -íne 'is/are.'

Still, to get éntone, the pattern of some deictic element plus accusative presumably was already possible in the Lefkas dialect. A likely candidate for this deictic element is the é already attested for Ionian islands in that very dialect, and

this, then, solves the problem noted above, in that it provides testimony, albeit of an indirect sort, for the unattested *é ton. It may therefore be concluded that Lefkas é tos could well have arisen in exactly the same way as ná tos, found elsewhere in Greek, *i.e.,* via an analogy from an accusative pattern *é ton. The analogical creation of a nominative weak pronoun posited for ná tos thus finds a parallel within Greek itself.

Furthermore, as suggested above, a similar sort of development can be seen in Hittite as well. While weak accusative and other oblique pronominal forms are found in most Indo-European languages, only Hittite, among the anciently attested languages,[11] has weak (enclitic) nominative forms, *e.g.,* COMMON.SG **as**, COMMON.PL **e** (also **at** or **as**), etc., which are used primarily (*cf.* Garrett 1994) with (unaccusative[12]) intransitive verbs. Moreover, according to Garrett (1990, 1994), they are a Hittite innovation and the process by which they arose was analogical. In particular, the model of the Hittite **a-**stem (Proto-Indo-European *o-stem) accusative (from *-om) to the **a-**stem NOMINATIVE (from *-os) seems to have been extended into the weak pronouns. Such an analogical extension is reasonable to posit since the accusative weak pronouns were presumably inherited into Hittite — they are found in some form in various ancient Indo-European languages, after all — and they pattern in form after the *o-stems (*e.g.,* **atta-** 'father'); an analogical schema as in (8) was possible, using **atta-** as a representative of the Hittite **a-**stems:

(8)	attan :	attas ::	an :		X,	X —> as
	NOUN/ACC	NOUN/NOM	he/WEAK.ACC		he/WEAK.NOM	

This analogy is parallel to what is posited above for Greek **tos**, the only difference being that Garrett looks to just a formal analogy within the nominal system alone, not tied to any one predicate as is posited with **ná** for Greek. This, however, is due to the fact that the type of predicate with which these innovative forms actually occur differs between Hittite and Modern Greek; they are restricted in Hittite, to unaccusative verbs only, but not as severely as in Greek and thus with a broader distribution. Still, for Garrett, their syntax is tied to his analysis of the organization of verbal arguments more generally in pre-Hittite, and the formal parallel to the creation of Greek **tos** is what is interesting here.

4. Answer to the Puzzle: Why Not in Any Other Balkan Language?

We are now in a position to understand why Greek **tos** is unique among the Balkan languages, despite the fact that the process by which it arose is natural — witness the fact that it is has parallels elsewhere — and that weak pronouns are found throughout the Balkan languages. In order for this analogical creation of **tos** to take place, four elements must be present. On the formal (morphological) side, there must be an obvious relationship in form between the strong and weak accusative (as in Greek, with **aftón/ton**), an obvious relationship in form between strong accusative and strong nominative (as in Greek, with **aftón/aftós**), and an obvious relationship in form between these two pairs of pronominal forms (as in Greek, with **aftón** figuring in both pairs). Further, on the syntactic side, what is needed is a context in which either nominative or accusative can be used interchangeably (as in Greek, with **ná** and neuter nouns). These ingredients taken together are what made the analogical emergence of the weak nominative **tos** possible.

As indicated by the parenthetical remarks above, all of these elements are present in Greek. However, in no other Balkan language does the system of third person personal pronominals have these four formal ingredients throughout all of the system; the relevant paradigms are given below for the Serbo-Croatian, Bulgarian, Macedonian, Albanian, and Romanian pronominal systems (where just the NOM and the cases with strong *vs.* weak opposition are listed),[13] with commentary as needed; moreover, the final ingredient is not to be found systematically in any of the languages either.

In Serbo-Croatian, for instance, as indicated in (9), there is a somewhat regular pattern formally in the opposition of strong *vs.* weak, specifically, [**nj(e)** + Weak] = Strong. Still, the form of the nominatives is sufficiently different from that of the oblique cases — in particular there is no initial **nj-** — that there is no basis for a proportional analogy that would take in nominative and oblique forms:

(9) Serbo-Croatian 3rd Person Pronominal System (Browne 1993)

Case/Number/Gdr	Strong	Weak
Nom.Sg.m	on	------
Nom.Sg.n	ono	------

Nom.Sg.f	ona	------
Acc.Sg.m	njega	ga
Acc.Sg.n	njega	ga
Acc.Sg.f	nju	je, ju
Gen.Sg.m	njega	ga
Gen.Sg.n	njega	ga
Gen.Sg.f	nju	je
Dat.Sg.m	njemu	mu
Dat.Sg.m	njemu	mu
Dat.Sg.f	njoj	joj
Nom.Pl.m	oni	-----
Nom.Pl.n	ona	-----
Nom.Pl.f	one	-----
Acc/Gen.Pl.m	njih	ih
Acc/Gen.Pl.n	njih	ih
Acc/Gen.Pl.f	njih	ih
Dat.Pl.m	njima	im
Dat.Pl.n	njima	im
Dat.Pl.f	njima	im

The same is true of Bulgarian and Macedonian third-person pronouns; no regular pattern linking the nominative and oblique is to be found, as shown in (10) and (11), respectively; note also that the regularities of form between strong and weak forms are more limited than in Serbo-Croatian:

(10)   Bulgarian 3rd Person Pronominal System (Scatton 1993)

Case/Number/Gdr	Strong	Weak
Nom.Sg.m	toj	------
Nom.Sg.n	to	------
Nom.Sg.f	tja	------
Acc.Sg.m	nego	go
Acc.Sg.n	nego	go
Acc.Sg.f	neja	ja
Dat.Sg.m	nemu	mu
Dat.Sg.n	nemu	mu
Dat.Sg.f	nej	i
Nom.Pl.m/n/f	te	-----
Acc.Pl.m/n/f	tjah	gi
Dat.Pl.m/n/f	(tjam)/im	
	na tjah	

(11)   Macedonian 3rd Person Pronominal System (Friedman 1993)

Case/Number/Gdr	Strong	Weak
Nom.Sg.m	toj	------
Nom.Sg.n	toa	------
Nom.Sg.f	taa	------
Acc.Sg.m	nego	go
Acc.Sg.n	nego	go
Acc.Sg.f	nea	ja
Dat.Sg.m	nemu	mu
Dat.Sg.n	nemu	mu
Dat.Sg.f	nejze	i

OF ALL THE SLAVS MY FAVORITES

Nom.Pl.m/n/f	tie	-----
Acc.Pl.m/n/f	niv	gi
Dat.Pl.m/n/f	nim/	im
	na niv	

The relevant forms are given in (12) for Albanian. These forms are actually demonstrative pronouns, pressed into service as the functional equivalent of third-person personal pronouns (Newmark *et al.* 1982:261-63); thus there is a distinction in "remote" *vs.* "near" deixis, roughly 'that' *vs.* 'this' (both forms are given here, remote first followed by near; variants in the dative-plural forms are given in parentheses):

(12)   Albanian 3rd Person Pronominal System
       (Newmark *et al.* 1982:261-63)

Case/Number/Gdr	*Strong*	*Weak*
Nom.Sg.m	ai/ky	------
Nom.Sg.f	ajo/kjo	------
Acc.Sg.m	(a)të/ këtë	e
Acc.Sg.f	(a)të/këtë	e
Dat.Sg.m	atij/këtij	i
Dat.Sg.f	asaj/kësaj	i
Nom.Pl.m	ata/këta	-----
Nom.Pl.f	ato/këto	-----
Acc.Pl.m	(a)ta/këta	i
Acc.Pl.f	(a)to/këto	i
Dat.Pl.m	atyre/këtyre	u (i)
Dat.Pl.f	atyre/këtyre	u (i)

Here too there is no obvious formal relationship between the strong and weak forms in the oblique cases that could serve as the basis for the extension of the opposition into the nominative and the creation of a weak nominative form. The same is true of Romanian, as shown by (13), where the variants in weak forms are conditioned

variants based on position and combinatorics and thus are essentially irrelevant to the concerns here:

(13) Romanian 3rd Person Pronominal System
      (Pop and Moldovan 1997)

Case/Number/Gdr	Strong	Weak
Nom.Sg.m	el	------
Nom.Sg.f	ea	------
Acc.Sg.m	el	îl/l
Acc.Sg.f	ea	o
Gen/Dat.Sg.m	lui	îi/i
Gen/Dat.Sg.f	ei	îi/i
Nom.Pl.m	ei	-----
Nom.Pl.f	ele	-----
Acc.Pl.m	ei	îi/i
Acc.Pl.f	ele	le
Gen/Dat.Pl.m	lor	le/li
Gen/Dat.Pl.f	lor	le/li

What these facts about the third-person pronominal systems show can be summarized as follows. There are some regularities holding among pronominal forms in the various Balkan languages, *e.g.,* South Slavic has a strong-weak form relationship in the obliques of [**n(j)e**- + weak forms] = strong forms, but there is no parallel relationship between accusative and nominative strong forms to round out a potential proportional analogy such as is found in Greek. Romanian perhaps comes close to having the necessary ingredients, at least for the nominative and most of the accusative forms, but the accusative-singular feminine form does not fit into any pattern. And, in any case, missing from Romanian is a more thorough connection of strong and weak forms in the genitive/dative case, a property which Greek does have, in, *e.g.,* the strong **aftú** 'his' and its weak counterpart **tu** 'his.'[14] Also, none of the languages except Greek has a syntactic context in which nominative and accusative are interchangeable so that an analogy could be activated.[15]

OF ALL THE SLAVS MY FAVORITES

In this regard, one interesting development in Bulgarian is worth noting, for it shows some movement in a direction that is potentially relevant, but the overall system still lacks all the necessary ingredients for the emergence of a weak nominative, in the account given here. That is, in addition to the accusative weak pronoun used with a deictic predicate such as **eto**, as in (3b), the construction **ka`de go** for 'Where is he?' also occurs, consisting of the locative question word plus a weak accusative; moreover, the casual speech (and somewhat impolite) form **gde** 'where?' can also occur here, thus **gde go** 'Where is he?' This usage is presumably modeled on **eto go**, much as Greek **pún don**, with accusative after **pún**, is based on **pún dos** (see *Section 3* above). Since Bulgarian has **ka`de e toj** as a more formal means of saying 'Where is he?' with the fuller form of 'where,' the 3SG copula **e** and the nominative pronoun **toj**, there is roughly a context in which nominative and accusative are functionally interchangeable, even if not exactly so (inasmuch as both *ka`de toj**, without a copula, and *gde toj**, with the reduced question word and a nominative, are impossible). Still, without an exact functional match but especially without the other necessary ingredients of form, it is unlikely that any further developments leading to a weak nominative form, such as a putative *goj, will take place.

5. Conclusion

What the preceding demonstrates is that the uniqueness within the Balkans of the Greek innovation leading to weak nominative forms like **tos** receives a principled explanation if analogy is taken to be at the heart of their creation. Both formal (morphological) and syntactic criteria need to be met for the conditions that could lead to the relevant analogy to be "activated," so to speak. If the creation of **tos** within Greek were just a matter of phonological reduction, then similar sorts of reductions might well be expected in at least some of the other languages, *e.g.,* giving **ta in Albanian from **ata** or **el or **le in Romanian from **ele**. To the extent, therefore, that such developments have not occurred, the pan-Balkan situation can be taken to provide another argument against taking **tos** to be merely a phonological reduction of **a(f)tós**; one must instead look more deeply, and bring morphology and syntax into the picture to account for **tos**.

Notes

* A version of this paper was read at the 12th Conference on Balkan Languages and Linguistics held at the University of Kansas in May, 2000. My thanks to Victor Friedman and Cynthia Vakareliyska for their very useful comments on the paper. I am indebted as well to Bob Dixon and Sasha Aikhenvald of the Research Centre for Linguistic Typology of La Trobe University (Australia) for providing me with the opportunity to finish this paper, through a generous Visiting Fellowship in the summer of 2001 at their Centre. An adapted version forms Chapter 8 of Joseph (Forthcoming).

1. Dixon 2000, however, has claimed that the Amazonian language Jarawara has no such elements at all.

2. My use of quotation marks with these terms is deliberate, by way of indicating just how slippery these elements are and of suggesting that they may not be the most illuminating of labels. Rather, I follow Zwicky 1985, 1994 in taking "clitic" not as the name of a distinct category that is a basic building block, an "atom" as it were, of the grammar, but rather as a cover term for elements with some word-like and some affix-like behavior but which nonetheless can be subsumed under the rubric either of words or of affixes. See Joseph 2000a, 2000b for discussion and specific application of this approach to the determination of "wordhood" in Modern Greek.

3. This is the term I prefer over the much vaguer and much abused and misused term "clitic" that is encountered so often in the literature; see footnote 2 for some relevant discussion and references.

4. There are occasional and sporadic examples of **tos** occurring after **íne** 'is/are' in early Modern Greek, as some examples in Roussel (1922:122) and Mohay 1998 show, suggesting a slightly wider range of occurrence for these forms beyond what is seen with **ná** and **pún**. Presumably, these represent extensions based on the use of **tos** with **pún**, which, as noted in *Section 3*, contains a reduced form of **íne**; thus, at a time when the **-n** of **pún** was still recognizable as belonging to the verb 'be,' **tos** must have been generalized into use with some fuller forms of **íne** not necessarily preceded by **pú**. Since these occurrences are not at all systematic, and can be made sense of in terms of the scenario posited here for the origin of **tos**, they are not considered further.

5.   The voicing of **tos** to **dos** in (5b) is a trivial change that occurs routinely after a nasal in Greek.

6.   Matras (To appear) discusses some very interesting facts from Romani dialects both within and outside of the Balkans concerning weak subject pronouns. In particular, he notes the existence of third-person subject "clitics," based on a stem **l-**, and considers these to be an old feature of Romani. Interestingly, in various dialects, these are restricted in use to presentatives, existentials, and interrogatives, not unlike the restriction on the distribution of the Greek forms in question here, raising a natural question as to the relation of the Greek and the Romani phenomena. However, since the dialects with these restricted forms include some in the Near East, it seems that they most likely arose within Romani in a pre-European stage. Thus Greek would still be alone in having

innovated such forms within the Balkans. Moreover, even if contact with Romani is involved in the Greek innovation, one still can ask why no other Balkan language reacted in the same way to contact with Romani, so the issues — and solution — discussed herein still seem valid. I thank Victor Friedman for bringing these facts to my attention and especially Dr. Matras for graciously sharing and discussing them with me.

7. As Victor Friedman has reminded me, since deictic constructions take nominative case in East and West Slavic, the use of the accusative in South Slavic may well itself be an innovation (though, given the semantics of deixis, as noted below, reinterpretation of a direct object with a deictic predicate as a subject is in principle possible for any language, so the accusative construction could be old within Slavic and the East/West nominative patterns could be the innovation). One can of course think of Greek influence as playing some role in the accusative construction of South Slavic deictics, a possibility which would most likely tip the balance in the controversy over the etymology of **ná** in favor of that proposed by Hatzidakis.

8. Although there was a general process by which unaccented word-initial vowels were lost between Medieval and Modern Greek, this loss was most regular with initial **e-** and **o-** and quite sporadic with a-. Moreover, it seems best to treat this vowel-loss as the result of resegmenting and/or contraction in connected speech (*e.g.*, earlier Greek **ero:to:** 'I ask' —> later **rotó**, via a path such as **ton ero:to:** 'him I-ask' —> **tone rotó**), a process that would not have been available with the sequence ***en atos**.

9. What I give here as **éntos**, transliterating what Pernot gives in the Greek alphabet, is presumably phonetically [é(n)dos] (with the post-nasal voicing seen in **pún dos**; see footnote 5), though Pernot is not specific on this point; nothing critical for my account hinges on this, however.

10. As noted in footnote 4, the occasional examples of **tos** with **íne** outside of the deictic and interrogative-locative use are likely to be extensions from those uses and thus not suitable starting points for **éntos** even if **én** is from the verb 'be.' Nor, as argued in footnote 8, is phonetic reduction a good possibility.

11. This qualification is needed since of course Modern Greek has the same type of element!

12. Unaccusative predicates are intransitives that govern as their single argument a nominal that is an object in terms of its grammatical (and semantic) relations. An example of an unaccusative predicate is *open* as in *The door opened*, since *door* does not have the thematic role properties typical of (underlying) subjects; rather, it has the thematic role of patient, thus one typical of (underlying) objects (as with *door* in *I opened the door*). This contrasts with intransitives that have a subject as their single argument, *e.g., run* as in *Robin runs daily*.

13. Neither are these conditions to be found in Romani, it seems, to judge, for example, from the third-person singular pronominal forms in the dialect of Romani in Agia Varvara in Greece (near

Athens). The information in Messing 1988 suggests that this dialect has, for instance, a third-singular masculine strong accusative form **kales** and a weak accusative form **les**, but a strong nominative form **vov** 'he,' so that no pattern quite like the Greek one occurs.

14. Note that even the extended strong forms like **aftunú**, mentioned above, have a somewhat regular morphological relation with the weak form, consisting of the weak form surrounded by **af-** as a prefix, **-n-** as a suffix, and a copy of what follows the **t-** of the weak form.

15. Admittedly, Hittite does not have that, and yet an analogy like the Greek one seems to have occurred. However, the analogy giving **as** in Hittite came at a point its development, in Garrett's account, when nominative and accusative were being sorted out into different functions, due to a shift the language was undergoing away from an ergative case-marking system.

References

Alexander, R. 1993. "Remarks on the Evolution of South Slavic Prosodic Systems," *American Contributions to the Tenth International Congress of Slavists* 1, Columbus, OH: Slavica Publishing Co., pp. 7-34.

_____. 1999. "Word Order and Prosody in Balkan Slavic Dialects, the Case of Thrace,". *Dialektologija i lingvisti™na geografija*, Sofia, pp. 61-73.

Andriotis, N. 1983. *Etimolojikó leksikó tis kinís neoelinikís*. Thessaloniki: Aristotle University of Thessaloniki (Institute Manoli Triandafilidi).

Banfi, E. 1985. *Linguistica balcanica*. Bologna: Zanichelli.

Browne, W. 1993. "Serbo-Croat," *The Slavonic Languages*, Comrie, B., and G. Corbett (eds), London & New York: Routledge, pp. 306-87.

Browning, R. 1983. *Medieval and Modern Greek*. Cambridge: Cambridge University Press.

Dixon, R.M.W. 2000. "Word in Jarawara," Paper presented at the International Workshop on the Status of "Word," Research Centre for Linguistic Typology, La Trobe, August 2000 (to appear as "The Eclectic Morphology of Jarawara, and the Status of Word," *Word: A Cross-Linguistic Typology*, Dixon, R.M.W., and A. Aikhenvald (eds), Cambridge: Cambridge University Press (2002)).

Friedman, V. 1985. "Balkan Romani Modality and Other Balkan Languages," *Folia Slavica* 7:3, pp. 381-89.

_____. 1993. "Macedonian," *The Slavonic Languages*, Comrie, B., and G. Corbett (eds), London & New York: Routledge, pp. 249-305.

Garrett, A. 1990. "Hittite Enclitic Subjects and Transitive Verbs," *Journal of Cuneiform Studies* 42, pp. 227-42.

_____. 1994. "Wackernagel's Law and Unaccusativity in Hittite," *Approaching Second. Second Position Clitics and Related Phenomena*, Halpern, A., and A. Zwicky (eds), Stanford: CSLI Publications, pp. 85-133.

Haiman, J. 1991. "From V/2 to Subject Clitics: Evidence from Northern Italian," *Approaches to Grammaticalization. Volume II. Focus on Types of Grammatical Markers*, Traugott, E., and B. Heine (eds), Amsterdam: John Benjamins, pp. 135-57.

Horrocks, G. 1997. *Greek. A History of the Language and Its Speakers*. London & New York: Longman.

Joseph, B. 1981. "On the Synchrony and Diachrony of Modern Greek *ná*," *Byzantine and Modern Greek Studies* 7, pp. 139-54.

_____. 1994. "On Weak Subjects and Pro-Drop in Greek," *Themes in Greek Linguistics. Proceedings of the First International Conference on Greek Linguistics*, Philippaki-Warburton, I., *et al.* (eds), Amsterdam: John Benjamins.

_____. 2000a. "Towards a Characterization of 'Word' in Modern Greek (Or, 'Do the Greeks Have an "It" for "Word?"')," Paper presented at the International Workshop on the Status of "Word," Research Centre for Linguistic Typology, La Trobe, August 2000 (to appear as "The Word in Modern Greek," *Word: A Cross-Linguistic Typology*, Dixon, R.M.W., and A. Aikhenvald, Cambridge: Cambridge University Press (2002)).

_____. 2000b. "Dialect Evidence Bearing on the Definition of 'Word' in Greek," Paper presented at the International Conference on Greek Dialects and Linguistic Theory, University of Patras, October 2000 (to appear in conference proceedings volume, edited by A. Ralli, B. Joseph and M. Janse, University of Patras (2001)).

_____. 2001. "Is there Such a Thing as 'Grammaticalization?,'" *Language Sciences (Special Issue — Grammaticalization: A Critical Assessment* 23:2-3), Campbell, L. (ed.), pp. 163-86.

_____. Forthcoming. "The Modern Greek Weak Subject Pronoun toß — Its Origins and Implications for Language Change and Language Structure: A Study in Grammatical Change," *Innsbrucker Beiträge zur Sprachwissenschaft* series (2002).

Klagstadt, H. 1963. "Toward a Morpho-syntactic Treatment of the Balkan Linguistic Group," *American Contributions to the Fifth International Congress of Slavists, Vol. I*, The Hague: Mouton, pp. 179-89.

Matras, Y. To appear. *Romani: A Linguistic Introduction*. Cambridge: Cambridge University Press.

Mohay, A. 1998. "To mistirio tu *námaste*," *Revue des Études Néo-helléniques (Publication des annales de la faculté des lettres et sciences humaines d'Aix-en-Provence. Nouv. Ser.*).

Newmark, L., P. Prifti and P. Hubbard. 1982. *Standard Albanian*. Stanford: Stanford University Press.

Pernot, H. 1934. *Introduction à l'étude du dialecte tsakonien*. Paris: Société d'Éditions "Les Belles Lettres."

Pop, L., and V. Moldavan. 1997. *Grammaire du roumain. Romanian Grammar. Gramatica limbii române*. Cluj: Editura Echinox.

Roussel, L. 1922. *Grammaire descriptive du roméique littéraire*. Paris: Ancienne Librairie Fontemoing.

Scatton, E. 1993. "Bulgarian," *The Slavonic Languages*, Comrie, B., and G. Corbett (eds), London & New York: Routledge, pp. 188-248.

Schaller, H. 1975. *Die Balkansprachen. Eine Einführung in die Balkanphilologie*. Heidelberg: Winter.

Zwicky, A. 1985. "Clitics and Particles," *Language* 61, pp. 283-305.

_____. 1994. "What Is a Clitic?," *Clitics: A Comprehensive Bibliography 1892-1991*, Nevis, J., B. Joseph, D. Wanner and A. Zwicky (eds), Amsterdam: John Benjamins Publishing Co., pp. xii-xx.