Phase bifurcation — causes and consequences in Benue-Kwa (Niger-Congo)*

Victor Manfredi
African Studies Center, Boston University

Abstract
Benue-Kwa (BK), the main branch of the Niger-Congo family, is a non-parametric typological space (Keenan & Stabler 1994, 2003; Newmeyer 2004) created by interaction of the internal and external dimensions of the human language faculty (I- and E-language, Chomsky 1986). §1 reviews a timing shift in cyclic spellout—the derivation-internal branching to PF—triggered by external factors (Manfredi 2005a, 2009a). In the remnant area (BK1) the lowest spellout domain is TP, including finite inflection aligned on or after the predicate root, but in the innovating zone (BK2) the lowest domain contains only open-class predicatons, thus is no bigger than VP, leaving closed-class clausal superstructure to a separate, subsequent phase.

§2 considers mechanisms of bifurcation and the status of prosody as determinant or determined. The initial state, today represented collectively by BK1, allows at most binary lexical ‘tone’ (lexical discrimination of roots by pitch), whereas this contrast is at least ternary throughout BK2. Some BK1 languages (Edó, Western Igbo) maintained late spellout even after obligatory finite suffixation had eroded to zero; short of appealing to analogy with optional suffixes, the only explanation is externalist and accidental: acquisition in BK2 of early spellout was triggered by a ternary lexical pitch contrast, most likely pushed by phonation effects which are still transparent in Gbè (Stahlke 1971). Such an inference is ruled out by standard views of prosody as “phonology” (Nespor & Vogel 1986; Ladd 1996) and of phonology as “different” (Bromberger & Halle 1989), but is supported by syntagmatic/metrical analysis of ‘tone’ (Bamba 1992; Manfredi 1995a,b; Dilley 2005a,b) and also on statistical grounds both in BK and in Sinitic (Manfredi 2009b).

§3 claims that other cross-BK differences are incidental to phasing. These include, not just the pragmatic (narrow-semantic) contrasts discussed in §1, but also certain options of overt morphosyntax: *in-situ* variables, affixal anaphors and marked accusatives in BK1; modals, inverse copulas and marked nominatives in BK2. All these are standardly treated as independent parameters, but are arguably contingent on prosody as understood in §2.

1. Timing shift
An enduring question in generative grammar is the relation between the abstract sound-meaning correspondence defining an I-language (“internal/intensional language”, Chomsky 1986, 20) and patterns of “external evidence” (Kiparsky 1971) in domains like acquisition, diachrony and (neural or artificial) processing. Recent decades have seen a departure from the null hypothesis of “no distinction of epistemological category” (Chomsky 1980, 109) between these phenomena, to a reductionist thesis of cognitive complementarity, factoring out from the *faculté de langage* elements such as “the content of lexical items” (Koster 1988, 5, cf. Borer 1984) and “principles of efficient computation” (Chomsky 2004, 105).

Complementarity predicts the interaction of internal and external properties of natural language to yield “empirical invariants” across “non-trivially different grammars” (Keenan & Stabler 2003). A case in point is an apparently abrupt shift in the grammar of Benue-Kwa (BK), the majority subgroup of the largest language family on the oldest inhabited continent. In BK, a patently external process of phonetic erosion opened the door to an internal timing shift from TP to VP as the phase spellout domain (cf. Chomsky 2001). This prosodic bifurcation of BK (§2) has a list of consequences (§3) for which no account seems possible, either from a strictly externalist, functionalist perspective (Givón 2009) or by appealing to revaluation of internal parameters of isomorphism based on precompiled triggers (Baker 1985, 1996; Kayne 2005).

1.1 Historical discontinuity
Following Westermann (1927), Greenberg (1963, 1972, 1974) established the Kwa and Benue-Congo subgroups of Niger-Congo (1a) and integrated “Bantu” into the latter over Guthrie’s (1962a,b; 1971) protests. He pregnantly added however that “Kwa and Benue-Congo are particularly close to each other and in fact legitimate doubts arise concerning the validity of the division between them” (1963, 39 fn. 13). Elugbe & Williamson concurred “that, pending the production of new types of evidence, Benue-Congo and Kwa form a single subgroup of Niger-Congo” (1977, 351). An alternate taxonomy (1b), floated on more and better wordcounts (Bennet & Sterk 1977; Williamson 1989), was subsequently abandoned, leaving the default picture of Benue-Kwa (or “East Volta-Congo” in homage to Stewart 1976) as a “dialect continuum” (Williamson & Blench 2000, 17f.).

Despite the failure of lexical evidence to detect any internal boundary above the level of individual clusters, BK ranges between extreme isolating and agglutinative morphosyntactic types (Williamson 1985; Hyman 2004), moreover the spectrum is not smooth as far as certain properties are concerned. Instead, a break point (1c) is defined concurrently on four dimensions, two audible (2a-b) and two interpretive (2c-d). As defined, these properties hold everywhere in BK2 (comprising the Gbè, Yorùbá-Igídà, Nupe-Ebirà and Iđomá clusters) and nowhere in BK1 (Kru, Ákan, Edò, Igbo, Cross, Plateau and Bantoid). On external grounds, the contiguous BK2 zone must be regarded as innovative, leaving BK1 as a conservative remnant. All four descriptions are concisely illustrated in a constructed paradigm contrasting Yorùbá and Igbo respectively in (3) and (4).
The historical picture above demands some modification of standard views of language change. On the one hand, it can be inferred that BK2’s emergence was not a matter of external drift alone, without the crucial intervention of I-language mechanisms, because properties (2c-d) are unlearnable from primary language data, thus they had to be side effects of (one of) the audible cues (2a-b) somehow or other. Conversely, however, the whole set in (2) can’t reduce to a single, omnibus-style I-language parameter, because all four of the BK2 values don’t necessarily correlate in other language families, for example (2c) holds in many southern “dialects of Italy” but (2a) doesn’t.

Nor can the BK2 blend in (2) be dismissed as an exotic coincidence: it holds for too many languages and speakers, given the absence of mixed settings. The only remaining possibility is that BK2 is the outcome of the external initial conditions of BK1 plus some internal mechanism. As it happens, there’s massive external evidence of suffix erosion, consistent with (2b). Starting from this fact, the BK bifurcation hypothesis (Manfredi 2009a) claims that all of the observed effects are consistent with a single shift in the domain of spellout, from relatively late (complement of C = TP) in BK1 to relatively early (bare VP) in BK2, assuming derivational phases (Chomsky 2001) and consistent with the finding in other language families that the timing of spellout can vary across closely related varieties (Barbiers 2008).

The inaudible traits in (2), addressed in previous work, can be derived from late spellout, very briefly, as follows. For (2c), if tense in BK2 necessarily fills the Aux position glossed FIN in (3) above, and if semantic and phonetic phases run in parallel (Chomsky 2004, 107), then BK2 languages with minimal finite inflection get ‘two bites at the apple’ of temporal reference via pragmatic default: a bare’ eventive VP receives a past denotation (Déchaîne 1991 following Welmers & Welmers 1968, 76), but additionally null T can anchor to topic time (Enq 1987), and past plus topic time straightforwardly yields present perfect. For (2d), the generalization is that “a sequence of aspectually unrelated events cannot be expressed in a single clause… unless each root is either local to Tense or audibly tensemarked (Manfredi 2005a, 13). In BK2 the second root fails this condition, hence its temporal reference is parasitic on the event structure of the first root, failing which it is uninterpretable.

Two more questions: how the timing of spellout was affected by the arity of lexical pitch contrast (§2), and whether the phase analysis is consistent with other known properties of BK grammars (§3). Provisional answers can be given to both.
Continuing to assume that the cue for early spellout was suffix erosion (2b), it remains to say why this would lead to ternary lexical pitch contrast (2a). A causal link between ‘tone’ and phrasing is unthinkable in modern phonology, which inherited from Pike (1948) a taxonomic view of lexical tones as paradigmatic primes, on a par with classical segments. This behaviorist idea repackaged in generative terms makes acquisition of lexical ‘tone’ look like formal complication—something infants as little economists should resist. Goldsmith writes: “The process of deautosegmentalization, as a task for the language-learner... consists of learning which sets of feature-specifications on separate tiers may be merged together” (1976, 164). But a dissident minority holds that ‘tone’ is irreducibly syntagmatic, whether as a matter of synchronic grammar (Sybesma 1989; Bamba 1991; Purnell 1998; Akinlabí & Liberman 2001), acoustics (Liberman & et al. 1993; Dilley 2005a,b) or acquisition (Harrison 1998ff.). Accordingly, pitch contrasts in roots aren’t forced on the child by word learning, but arise spontaneously unless blocked for some reason. And even when blocked in roots, pitch contrasts can develop in lexical contexts with more structure, e.g. zero derivations like English export, exportation.

It is relevant to observe that suffix erosion by itself was not a sufficient condition for phase bifurcation, thus some BK languages display suffixless finite verbs and yet remain fully BK1. In Edó, a putative ‘suffix’ does appear in certain finite contexts such as an object relative clause (5b), but this item is demonstrably epenthetic with a phrase-final and prosodic distribution, as shown in (6). The generalization is that Edó finite inflection is primarily realized on the root, as in the minimal contrast between past/finite (5a) and nonpast (5b), but just in phrase final position the finite pitch accent is realized as a full trochaic foot [HL] as in (5c) and (6c), cf. (7). There’s no epenthetic suffix on the phrase-final root in (5d) because there is no pitch accent (the sentence is static, effectively nonfinite, even though the root is eventive), and there’s no epenthetic suffix on the finite root in (6d) because it is not phrase-final, thanks to the phrase-final adverbial.

Western Igbo (9) has no reflex of the finite suffix of Ónicha (8) and other eastern varieties, but maintains late spellout in an impressive variety of ways: either by deaccenting the root, which in tonological jargon is lexically H but pronounced L in (9a-b) no less than in (8), or else by deaccenting the subject clitic on the VP’s edge just in case the root is accented (lexically H), cf. (9c) vs. (9d) where the root is lexically L/unaccented. A mix of both strategies also occurs (9b).

2. Tones and times

_Hoenigswald is right in asserting that the loss of suffixes through sound change is a well known and frequent event... But on the other hand, it is well to recall... that a mere sound change is not enough for a grammatical upset._ (Jakobson 1949/1971, 112)
Given all these ways to save late spellout, what could dissuade the learner of any BK language from phasing T and VP together? Both (8) and (9) exploit the existence in Igbo of binary lexical pitch, whether by effacing this information on the open-class root or by mapping it onto the closed-class (clitic) subject. Accordingly, the exceptionlessly ternary contrast of lexical pitch in BK2 suggests that this fact in itself can block any audible dependency between T and VP. E-language pressures towards a ternary contrast certainly have occurred in BK2, as shown by phonation effects (Halle & Stevens 1971) which are synchronically productive in very many Gbé varieties (Stahlke 1971; Gbêto 2002). That in itself would trigger early spellout, so long as the structural preservation of ternary lexical ‘tone’ contrast forces the left edge of VP to begin a foot. Such a notion is nonsensical in autosegmental terms, but ample supporting evidence appears in Yorùbá (BK2). As noted in (10) by the contrast between vertical and diagonal constituency lines, Yorùbá L is represented as an adjunct to H, the head of the foot. Some structure of this type is required in order to account for the syntagmatic phonetic behavior of Yorùbá L, and it arguably entails that a Yorùbá root with lexical L is foot-initial.17

Historically therefore, once segmental suffixes had eroded, another (probably related) external development (ternary pitch on lexical roots) was sufficient to reset an internal syntactic parameter (late—early spellout), with many consequences.

3. Prosodically-driven syntax

The notion that prosody can cue syntax is not unknown in acquisition literature (Donati & Nespor 2003), though standard grammatical architecture seems to rule out such an implication, i.e. if prosody is “phonology” (Ladd 1996) and phonology is “different” (Bromberger & Halle 1989). For example, Abo [h] (2007) points to a striking outcome of the developments sketched above: in BK2, logical variables like narrow focus are unable to occur to the right of the finite verb e.g. (11a), but in BK1 this possibility is generally available e.g. (12a). This contrast extends to what can only be described as the inverse copula of BK2, e.g. (11b-c), versus its canonical counterpart of BK1, e.g. (12b).18

Although the leftward displacement of variables is most commonly treated as wh-movement, there are compelling reasons to do otherwise, such as the failure of weak crossover effects (13). A possible clue as to what, if not feature checking, forces Yorùbá narrow focus left, can be inferred from the failure of tonal footing in the complement of a transitive verb. A Yorùbá verb root cannot bear L tone before an overt complement of argument type; this is resolved by nonparsing the lexical L (14a), yielding nonlexical M, the pronunciation of tonelessness (Akinlabí 1985). This effect is absent before an adjunct (14b) or an empty category (14c), posing a problem for narrow phonology (Carstens 1987).

The nonparsing of the lexical L in ‘sell’ in (14a) follows from the foot structure in (10), because a verb can’t adjoin to its own complement (Déchaine 2001), and the failure of postverb narrow focus could be similar. The affirmative inverse copula ni is in complementary distribution with affirmative finite inflection HFIN as well as with nominative/affirmative
subject clitics such as mo ‘1S.NOM’, all of which realise the Sigma (affirmative/negative) archi-feature in T (Laka 1990; Déchaïne 1992, 1995). If so, then left peripheral focus as in (13) and (14c) has two doses of finite affirmation, but this is not biclausal: the bracketed remnants in these examples are not well-formed relative clauses. If Yorùbá VP is a spellout domain, then any internal argument is already linearized before merger of the affirmative focus operator Sigma in T, therefore no focus interpretation is possible unless this operator is re-merged as a “proxy category” (Nash & Rouveret 1997) above T, and the argument re-linearized—an example of “scopophobia” (Manfredi 1997, 2006) or “foot-driven movement” (Craenenbroeck 2006), neither of which appeal to the circular device of feature checking/attraction.\(^\text{20}\)

Another morphosyntactic difference that seems to track spellout domains is the form of reciprocals. In BK1 these can be expressed in between the predicate root and finite aspect. In Êgbo this option is restricted to adjuncts (15a), whereas Kirundi does so also with arguments (15a) but the contrast with BK2 languages remains: in Yorùbá, corresponding sentences of both types require nonclitic (XP) expressions containing a possessed dummy noun such as ‘body’ (16).\(^\text{21}\)

\[\text{Êgbo (BK1, Úchêchâkôwù 2003)}\]

(15)a. Êgigè na ìge e-kwè-kọ-ri-ta-go. N. and I. FIN-agree-COMIT-APPL-BACK-ASP ‘N. and I. have agreed with each other’

Kirundi (BK1, Ndayiragije 1999, no tone)

b. Abo bagabo ba-a-sambuir-an-ye. those men 3P-T destroy-COMIT-ASP ‘Those men have destroyed people/each other’

In the above contrast, it might matter that T and VP are not phasemates in BK2. Reciprocals respect binding condition A: their antecedent must be found within the nearest cyclic domain, modulo intervention effects (17b). Of course a suffixal anaphor can’t exist in Yorùbá because Yorùbá has no suffixes at all, but why not? Yorùbá has a perfectly good comitative predicate (18a), but this can’t replace an argument (18b) à la Kirundi (15b). So apparently, the phase/cyclic boundary between VP and T in BK2 forces anaphors to be phrasal, for the sake of locality to T.

\[\text{English}\]

(17)a. They said [Jo and Bo saw [pictures of each other]].

b. *[Jo and Bo saw [Sue’s pictures of each other]].

\[\text{Yorùbá (BK2, ‘S. Oyèlãrùn, p.c.)}\]

(18)a. Êgigè àti ìgè-è jò lo. N. and I. FIN COMIT go ‘N. and I. accompanied each other/went together’

b. Êgigè àti ìgè-è jò pa *(èran). N. and I. FIN COMIT hit ‘N. and I. cooperated in killing an animal’

Two more consequences of phase bifurcation can be signalled here. BK1 is rich in auxiliaries but lacks morphosyntactic modals: as shown by Êménanjò (1985) for three Êgbo varieties, all auxes require finite inflection, either by deaccenting/loss of lexical H as in (19a), or by segmental affixation in an appropriate context. By contrast, Yorùbá possesses two series of auxes, one of which—the irrealis set—inflects not at all (Oyèlãrùn 1992; Déchaïne 1992, 1995), cf. (20a). Secondly, morphologically specialized accusative forms—whether full arguments (not shown here) or pronominal clitics—are found throughout BK1, but nowhere in BK2, cf. the contrast between (19b) and (20b). Thus in Yorùbá irrealis clauses, a clitic subject can’t be nominative e.g. mo ‘1s’, but must be a notional ‘accusative/genitive’ e.g. mi ~ n ‘1s’ (Manfredi 2003b).

\[\text{Êgbo (BK1)}\]

(19)a. ìgbọ à ã ìgbọ fù. vehicle this go.FIN ing-flying top.GEN ‘The plane is going to take off’

b. ò lá n’àn yà 3S.NOM look-FIN 3S ACC eye ‘S/he regarded her/him/it’

\[\text{Yorùbá (BK2)}\]

(20)a. òkò yóò fò. (n.b. *òkò-ò yóò fò) vehicle MOD jump ‘The plane will take off’

b. ò wò ò. 3S.FIN see 3S ‘S/he regarded her/him/it’

As to the absence of morphosyntactic (as opposed to logical-semantic) modals in BK1, it seems impossible for an item to be base-generated in T (Roberts & Roussou 2003) unless T occupies a separate phase from VP, assuming that internal merge is more economical than its external counterpart. The complement of any Êgbo aux is transparently nonfinite, as in (19a). As to specialized accusatives, this phenomenon is ascribed in many frameworks to ‘case competition’ with nominative at the point of spellout (e.g. Bittner & Hale 1995), but if T and VP occupy different spellout domains, a potential configuration of this kind is unavailable.
But [Sapir] then proceeded to raise a new question: do the phonemes he postulated have “psychological reality”? To answer this question he turned to other kinds of data, which is sometimes called “psychological evidence,” that is perceptual tests of various types. … But another question arises. Why didn’t the “linguistic evidence” suffice to establish “psychological reality”? … In fact, in this case the “linguistic evidence” may well be more persuasive than Sapir’s “psychological evidence”. … What the best evidence is depends on the state of the field. The best evidence may be produced by yet unexplained facts drawn from the language being studied, or from similar facts about other languages, or from psycholinguistic experiments, or from clinical studies of language disability, or from neurology, or from innumerable other sources. … But there is no distinction of epistemological category. In each case, we have evidence—good or bad, convincing or not—as to the truth of the theories we are constructing; or if one prefers, as to their “psychological reality” though this term is best abandoned as seriously misleading. … One might ask at this point, once again, what is the distinction between a theory held to be true of the speaker’s capacities and implicit knowledge, on the one hand, and a “psychological hypothesis” on the other. I see no distinction. … No matter how broadly we cast our net in seeking evidence, it will always be true that our theories leave open innumerable questions about mechanisms. This will remain true no matter how far an inquiry into language and its use proceeds. When should we be willing to say that we are presenting psychological hypotheses and describing “inner psychological mechanisms?” (1980, 108ff., original emphasis)

2. Newmeyer (2004), as well as recent re-evaluations of classical, interlistal pro-drop (Neeleman & Szendrői 2007), are relevant here.

3. The current Bantuist handbook admits that that it’s “impossible to draw a clear line between Bantu, however defined, and non-Bantu Niger–Congo” (Nurse & Philippson 2003, 3, cf. Marten 2006).

4. Carelessly or misleadingly, the current Bantuist handbook labels a family tree incorporating the division in (1b) as having been “adapted from Williamson & Blench (2000)” (Schaedeberg 2003, 154).

5. Caveats: the status of (2c) in Kru and Nupe, and that of (2d) in Idoma, is not clear from available sources.

6. Tonemarks: [ ] = high, [ ] = low. Conventions: in BK2, no mark = mid; in BK1, no mark = same as preceding mark.

7. The status of (2c) in Kru and Nupe remains to be determined.

8. Definition: A serial construction consists of multiple predicates of the same lexical category (e.g. ‘verb’) in one clause, thus sharing one grammatical subject and one tense (cf. Stahlke 1970, 60, 80; Déc’haine 1993, 201). The status of (2d) in Idoma is unknown to me.

9. This inference can be avoided by functionalist grammar allowing “diachronic drift along universal semantic rivers” (Givón 1975, 93) and also by model-theoretic philosophy countenancing “semantic parameters” (Chierchia 1998). For both alternatives, however, the apparently exceptionless correlation of (2c-d) with (2a-b) across the large BK population remains unexpected.

10. Assuming that (2c) describes the following data from Rolffs (1969: 48f. cited by Cordon 1997, 88). On the other hand, (2c) and (2d) may be positively correlated, insofar as Cardinaletti & Giusti (2001) observe at least marginal serialization in parts of southern Italy. The issue is hard to evade in Mandari, both because the number of lexical ‘tones’ (four) may not be independent of the minimally bimoric length of the nucleus in lexical roots (Duamnu 2000, 2004) and because the status of (2c) is controversial there (Paul 2004).

11. Calabria

(ii.a) [tor’nasti]

Salento

(b) [tir’asta]

12. For some reason, irrealsis/infinital interpretation of the second root is impossible both in BK2 and in “Atlantic creoles” (Mufwene & Dijkhoff 1989, 326 fn. 26), although the situation in Mandarin is apparently different (Paul 2004).

13. Traditional Africanists held that the eroded suffixes were originally tone-bearing, and their residual tones drifted leftward onto roots so as to create new tonemes, and the same has been said for Cantonese, but both claims are implausible: the proposed phonetic rules are ad hoc, not to mention that the eroded suffixes lacked intrinsic lexical pitch specification (Manfredi 2003a, 2009b).

14. Indeed, the so-called bare ‘noun’ in BK has many phrasal characteristics (Manfredi 2004, 2009b).

15. The trochaic character of the tonal foot in Edó is claimed by Manfredi (2003a, 2004). The paradigm in (6), kindly confirmed b. Q. Ogie (p.c.), was constructed from observations by Baker & Stewart (1997, 44) and Stewart (2001, 178f.). The so-called “Bantu verb-final vowel” (a) in complementary distribution with morphologically specified aspect (a) also diagnoses phrasal epenthesis. Thus across BK1 we see that alleged ‘suffixation’ is less syntactic than ‘tone’, contra Stewart (1997), Baker & Stewart (1999).

16. Descriptively, finite inflection can be a pitch accent in Edó, because all predicate roots are “toneless” thanks to a regular tradeoff between pitch and moraic/syllabic structure (Amayo 1975). The same item glossed noncommitally as -AFF(i) in Edó above, is further specified as -CL(itic) in eastern Igbo (8) due to interpretive effects (Manfredi 2005b) that may extend to Edó (Stewart 1997).

17. Relevant observations, mainly due to linguists working at the Universities of Ibadan, IFE and IGR in their heyday before Babangida fascism and IMF structural adjustment, are reviewed by Manfredi (1995a), Déc’haine (2001) and Akinlabí & Liberman (2001). An urgent project is to survey possible remnants of tone-splitting in Yorùbá, whose lexicon is the best-studied in Niger–Congo.


19. No tones in the source. A Yorùbá-like, left-peripheral paraphrase of (12b) is also possible: Atade ne ade a me òòe (Boudi 1974, 25).

20. An exception is focus of an intransitive, which can occur low: Kö lọ lọ lọ ’S/he didn’t go at all’ (Awóbìùjì 1972, 126).

21. In (16) and (17), both of the (b) forms are ambiguous, but crucially not in the same way, thus BK1/2 are non-isomorphic.
References

Aboh, O. [2007]. Leftward versus rightward focus; the Kwa-Bantu conspiracy. *SOAS WPL 15*, 81-104.


