LINGUISTIC FORM COMPRESSION IN DAGAARE

Adams Bodomo, *K. K. Luke*, and *Owen Nancarrow*

Department of Linguistics
University of Hong Kong
abbodomo@hku.hk
November 2001

Abstract

Linguistic form compression is a productive morphosemantic process in Dagaare, a Gur language spoken in northern Ghana and other parts of West Africa. Linguistic form compression results in 'short forms' of words, phrases, and sentences. A 'short form' is a word (or 'fixed expression') which is formed through the reduction in quantity of the morphophonological material of a longer word, phrase or sentence while keeping the meaning intact. This is a second-order coding in language in the sense that it is a process which takes an existing linguistic form as input and produces a compressed form as output while keeping the meaning of the original form intact. The article proposes some general principles governing this process. The article focuses on three types of data in Dagaare: (i) phonetic short form formation from final vowel elision, (ii) the compression of compounds and long words, and (iii), the compression of proverbs.

La compression des formes linguistiques est un processus morphosemantique très frequent en dagaare, une langue gur du nord du Ghana et d'autres lieux en Afrique de l'ouest. La compression des formes linguistiques produit des 'formes courtes' des mots et des phrases. Une forme courte est un mot qui est le resultat de la reduction de la quantité du materiel morphosemantiqare d'un mot ou d'une phrase longue alors que son sens reste intact. C'est un deuxieme ordre de codification en langue dans la mesure ou c'est un processus qui prend une forme existante comme donnée et produise une forme comprimée alors que le sens reste le même. Cet article propose des principes generaux qui sont a la base de ce processus. L'article porte sur trois types des donnees en dagaare: (i) la formation des formes phonetiques par l'elision des voyelles finales, (ii) la compression des mots longs et complexes, et (iii) la compression des proverbes.

0. Introduction

Short forms are phonologically and lexically reduced forms of longer words and phrases of the language. A salient feature of the short form is that it is possible to retrieve the full semantic referents/contents of the long form from the short form. The nature of the orthography of the language seems to present important consequences with regards to the principles governing the production of short forms from longer words. This is an important issue for those languages that have an established writing system. Chinese and English are languages that belong in this category (Luke and Nancarrow 1999).

In this paper, we examine a scenario of short forms in the absence of an established orthography. This is a situation in which some languages do not have an established orthography or one in which a larger segment of the speakers of the language are not literate in an established orthography, as is the case with the Dagaare language of Ghana, West Africa.

1. Phonetic short forms

1.1. Short forms from final vowel elision/apocope

Final vowel elision is a pervasive phonological process in Dagaare and other Gur languages (Rialland 1985 for Moba; Kennedy 1966, Bodomo 1997, and Anttila and Bodomo 1997 for Dagaare; Hyman 1993 for Dagbane; and Cahill 1999 for Konni). In Dagaare, there is the tendency for the final vowel which is often not a distinctive functional morpheme to be dropped in various speech situations, such as fast speech, recitations and songs. In (1) below the final high and mid vowels are elided, thus giving us reduced forms of the words in which they are found.

(1) a.   *mírì*   ⇒   *mír*   'rope'
          wírì   ⇒   *wír*   'horse'
          bíirí   ⇒   *bíir*   'children'
          píírì   ⇒   *píír*   'sheep'
          bóórí   ⇒   *bóór*   'goats'
          bírì   ⇒   *bĭr*   'seed'
          dàgáárì   ⇒   *dàgáár*   'Dagaare'
          zùmmú   ⇒   *zŭm*   'fish'
          lóngó   ⇒   *lóg*   'toad'
          sóŋó   ⇒   *sŏŋ*   'wood carrier'
          póg-ó   ⇒   *póg*   'woman'
          tègé   ⇒   *tèg*   'only'

     b.   *póg-bó*   ⇒   **pógb*   'women'

Sometimes, even vowels representing a distinctive functional morpheme, such as the singular suffix marker in the word for woman in (1b) can be elided, with the understanding that the root word also stands for the singular form of the noun. The final vowel of the item *póg-bó* cannot be elided probably because the consonant segment alone cannot indicate the plural form of the word, and such syllable structure in coda position is forbidden in Dagaare.

Such a pervasive case of final vowel elision gives rise to the possibility of reduced forms for the majority of words in the language that ends in vowels.

One interesting difference between Dagaare on the one hand and English on the other is that apocope (deletion of final element in the word) is more productive than aphaeresis (the deletion

of an initial sound in a word as seen with English '*He's come.*') and syncope (the deletion of sound in the middle of the word e.g. *police - p'lice*).

2. Compounds and long words

Phonetic reduction involving apocope of words ending in high front vowel and mid vowels is one type of process inducing short forms in Dagaare. Another type deals with lexical shortening and the linguistic domain for such a process is mainly compound words and other words of more than one syllable. These may be categorized into personal names, travel names, and common compound nouns.

2.1. Personal names

Every personal name in Dagaare culture has a meaning, however oblique it may be. The birth of a child is often an opportunity for the head of the family or any of the parents to come up with a concise statement, a view of the world, a maxim, a philosophical observation, or some political indirection in the form of a personal name for the new-born child.

As may be seen in (2), it could be an expression of joy of being the custodian of a new-born child in phrases and exclamations such as 'Joy reach me' or 'Good tidings!' It could be an expression of gratitude to God for getting a much awaited child, as is the case in (2c&d) in expressions such as 'I look up to God' or 'God is King'.

(2) a. *Nòɔ̀-tá-má* ⇒ *Nòɔ́*
         joy-reach-me   'Joy'
         'Joy'

     b. *Yèlvìlàá* ⇒ *Vìlàá*
         deed-good   'Good'
         'Good deed' = Good tidings

     c. *Ngméǹ-kà-ńg-kààrá* ⇒ *Ngméǹ*
         God-COMP-1SG-watch-IMP   'God'
         'I look up to God.'

     d. *Ngméǹ-lá-nàá* ⇒ *Ngméǹ*
         God-FOC-king   'God'
         'God is king.'

     e. *Sòńgléédòńg* ⇒ *Sòńg*
         good-become-enmity   'good'
         'My good deeds breed enmity.'

3

More often than not, this name given to the infant is quite long, and it is another feature of the naming situation to select some part of the original name to represent the name of the child. This is not consciously done and it is not part of the formal naming ceremony. It is often an informal, spontaneous outcome, beginning with closer relatives of the child and finally getting accepted by the community as an alternative way of referring to the child. Thus, short forms like *Ngméǹ*, *Nòɔ́* and *Vìlàá* are used by family and friends to refer to the person who has any of the above longer names.

Now, the question is, is there a systematic way in which this radical shortening process takes place? Which segments of the longer words are targets for the formation of shorter forms? From the above data, it may be observed that two position morphemes are targets for the short forms, namely the first position morphemes, *Ngméǹ* and *Nòɔ́* and the end position morpheme, *Vìlàá*. This observation seems to be congruent with the case in Chinese and English, in which any of the front or end morphemes are used to form the short forms.

What, then, is the principle guiding the choice of candidates for the short forms in Dagaare names? To answer this, let us look at more cases of names. We consider yet another type of names in Dagaare.

While the above involve names given to infants by their guardians, there is a different situation in which Dagaare speakers give themselves names. Travel names are a particularly interesting aspect of 'autonyms' among the Dagaaba. It is often the case that young Dagaaba travel from their homeland in the north of Ghana to the southern parts of the country in search of jobs and better standards of living. This is often seasonal. Upon reaching these parts they are often in identifiable Dagaare speaking communities. Autonymy has become a culture of these Dagaare-speaking communities in southern Ghana. Members give themselves new names while they are in the South of the country, and may even use these when they go back to Dagao, home of the Dagaaba. As with the personal names at birth, these travel names often encode some view of the world or some experiences of the person carrying the name. Let us consider a few of them in the next section.

2. 2. Travel names

(3) a. *Gàà-zíé-nyé* ⇒ *Gààzíé* ⇒ *Gàànyé*
         go-place-see   go-place   go-see
         'travel and see'   ' travel and see'

     b. *Yɔ́-bɔ̀ng-yáǹg* ⇒ *Yɔ́-bɔ̀ng*
         roam-know-knowledge   roam-know (knowledge clipped)
         ' Travelling gives us a better understanding of the world.'

     c. *Wòng-túó-tá* ⇒ * *Wòngtúó*
         suffer-pain-reach
         'enduring journey'

4

The above three names are typical of the structure of travel names. They often involve two or more verbs and some objects or adverbs of place, such as 'go-place-see', meaning travel and find out, 'roam-know-knowledge', meaning traveling gives one a better view of the world or even 'suffer-pain-reach', meaning the individual involved has experienced some painful encounters during the journey to the south of Ghana.

In (3a) and (3b), the short form comprises the first and middle morpheme, *Gàà zíé* 'go-place', and *Yɔ́-bɔ̀ng* 'roam-know' respectively. The prediction then would be that the short form of the third should also follow the previously observed pattern. However, this is not the case. How then does one predict the right candidate for the short form?

It seems that two principles work here to predict the right candidate for short forms. These may be stated as follows:

(4)   a.   Principle of semantic equivalence: the short form and the longer   counterpart of
            a name must have the same referent or semantic content {perspicuity}

       b.   Principle of salience: to form a short form, pick out the most semantically salient
            segment of the long form

Principle (4a) will rule out (3c) because it seems that *Wòngtúó*, the short form, only expresses the idea of suffering and does not adequately translate the idea of perseverance, ending in a satisfactory accomplishment of the act of travelling. In the same vein, if one were to select *tá má* in *Nòɔ̀tá má* or *Yèl* in *Yèlvìlàá* one may not be achieving a semantically perspicuous version of the original name. So the principle of semantic equivalence works to predict why certain morphemes by themselves would not be adequate candidates for the short form.

The principle of salience (4b) seems to be the principle that actually predicts or guides the selection of the winning candidate for the short form. Intuitively, the most compelling aspect of the this informal, spontaneous aspect of assigning short forms to names is that it is usually the most meaningful word in the phrase that gets retained as the short form. Thus we have *Ngméǹ* in any of the longer names containing this, *Nòɔ̀* in *Nòɔ̀tá má* and *Vìlàá* in *Yèlvìlàá*; *Gàà zíé* in *Gàà zíényé*, *Yɔ́bɔ̀ng* in *Yɔ́bɔ̀ngyáǹg*. The mere mention of the name of God alone shows the gratitude of the person or his recognition of God as author of good deeds, thus rendering all other items in the name less salient. In *Nòɔ̀tá má*, it is basically about the expression of joy and once the word for joy is mentioned the message is fairly well translated. In *Yèlvìlàá*, the primary target is to express the notion of good tidings and thus *Yèl* would be less salient in the circumstances.

In the case of travel names, the salience principle will ensure that the most important verbs are picked up, with or without the intervening objects or adverbials. In *Yɔ́bɔ̀ng* both verbs are picked up, and their meaning makes the object 'knowledge' quite redundant or less prominent. In the case of *Gàà zíényé*, two alternatives are available, *Gàà zíé* or *Gàànyé*. In the first alternative, the addition of the adverb of place *zíé* renders the second verb less necessary, while

if the second verb is chosen the adverb of place is not necessary. The first alternative seems to be more frequent. It seems then that the principle of salience selects the most salient verb *gàà*, and leaves the choice of the adverb of place or a second verb to some other principles or factors. The salience principle would also select *wòntúó* as the most prominent semantic items but the hedging acceptability of this candidate as the short form demonstrates a conflict between the principle of salience and the principle of semantic equity.

It thus seems possible to predict that any of three types of short forms can occur: clipping from the first, middle or final position of the input. However, this typology is constrained by principles of interpretation such as the salience principle and the principle of semantic equivalence. Dagaare seems to favour the first position candidates in this typology of short forms, as demonstrated by the prevalence of first position morphemes as short forms in the data. The middle position typology is the least preferable in Dagaare. However, this preference for first position morphemes as candidates can be overridden by stronger principles. Let us push the description further with more types of short forms in the language.

2.3 Common compound nouns

Compounding in Dagaare involves any combinations of lexical items from the same or different word classes such as N+N or N+A combinations. Some, but not all, compound words are also subject to the process of shortening in Dagaare. Two positions are targets for this shortening, the first and last members of the compound. In (5) – (8), I list about four compound words, indicating their short forms where available, and grammatical contexts in which both short and full forms may be used as the same grammatical functions.

(5)   a.   *dɔ́ktà-yírì*   ⇒   *dɔ́ktà*
            doctor-house   doctor
            'hospital'   'hospital'

        b.   *ǹ*   *gè-ré*   *lá*   *dɔ́ktà-yírì*
            1.SG go-IMP FOC doctor-house
            'I am going to the hospital.'

        c.   *ǹ*   *gè-ré*   *lá*   *dɔ́ktà*
            1.SG go-IMP FOC doctor
            'I am going to the hospital.'

(6)   a.   *kùr-wìrí*   ⇒   *wìrí*
            iron-horse   horse
            'bicycle'   'bicycle'

        b.   *kà*   *ń*   *dé*   *má*   *kúrí-wìrí*   *kúlí*   *yàà*
            COMP 1.SG take 1.SG.GEND.   iron-horse go.home PART
            'Let me take my bicycle home.'

6

    c.   *kà   ń   dé   má   wìrí   kúlí   yàà*
        COMP 1.SG take 1.SG.GEND. horse go.home PART
        'Let me take my bicycle home.' [but could also mean horse here]

(7)   a.   *dàlúgrí*   ⇒   *lúgrí*
        bar   bar

    b.   *à   dàlúgrí nyέ kɔɔ̀-ὲ   lá*
        DEF stick-bar this   break-PERF FOC
        'This bar is broken.'

    c.   *à   lúgrí nyέ kɔɔ̀-ὲ   lá*
        DEF bar   this break-PERF FOC
        'This bar is broken.'

(8)   a.   *sàsìέ*   ⇒   *sìέ*
        'wind'

    b.   *sàsὲέ fúúró   lá*
        wind   blowing FOC
        'It is windy'

    c.   *sὲέ   fúúró   lá*
        wind   blowing FOC
        'It is windy'

However, certain compound words do not easily lend themselves to shortening and this will need some amount of explanation.

(9)   a.   *kpànkpàngbíél-í* ⇒   **kpàngbíélí*
        arm-arm-pad-SG
        'elbow'

    b.   *gbéńgbéláá*   ⇒   **gbéláá*
        'dry place'

3. Proverbs

A third category of word classes that exhibits visible form compression is proverbs. As with many West African languages, Dagaare makes a pervasive use of proverbs as a form of verbal indirection. Proverbs are so much frequently used that many average speakers of the language can easily infer their meanings and application in the slightest of contexts and as soon as one begins to use them.

Ease and frequency of use and predictability of context are fertile environments that lend themselves to linguistic form compression. Proverbs seem to satisfy all of these, hence the productivity with which they lend themselves to shortening in Dagaare. In (10a) to (19a) below, I list about ten proverbs, and indicate a most likely candidate for shortening in (10b) to (19b) constructions. The constructions in (10c) to (19c) show contexts in which the short form would be used in the language.

(10)   a.   *tá   tá   má séé   lá   bàrè   má ǹ   nú*
                NEG reach me be.better TOP leave me my hand
                'Don't get near me is better than leave my hand.'
                ENG: Prevention is better than cure.

       b.   *tá   tá   má*
                NEG reach 1.SG
                'Don't get near me.'

       c.   *à   yélé nyé nàng tá   tá má yélé lá   yàà*
                DEF. matter these just NEG reach 1.SG matter FOC PART
                'This is a case of "Prevention is better than cure".'

(11)   a.   *tágè kà ń tágè lá   kyèɛ̀-né   ngmáné*
                pull   and I   pull TOP tear-IMP   calabash
                'If we fight over a useful thing we may both lose it.'

       b.   *tágè kà ń tágé*
                pull   and I   pull
                'If we fight over a useful thing'

       c.   *tágè kàń táge yélé lá à   ámè*
                pull   and   pull   matter FOC DEF these
                'This is a case of "fighting over a useful thing" '.

(12)   a.   *bì-bìl-tòng-zágr-á   bá   dì-ré   bón-káà*
                child-small-send-refuse-AG NEG   eat-IMP   thing-oil
                'A child who refuses to go on errands does not get the opportunity to eat juicy things'

       b.   *bì-bìl-tòng-zágr-á*
                child-small-send-refuse-AG
                'A child who refuses to go on errands'

       c.   *bì-bìl-tòng-zágr-á   lá*
                child-small-send-refuse-AG FOC
                'That is one who refuses to go on errands.'

(13) a. *ngmàà-dà-kóré bá yìnn-é tèɛ̀ dúó-ù*
monkey-male-old NEG forget-IMP tree climb-NOM
'An old monkey never forgets how to climb trees.'

b. *ngmàà-dà-kóré*
monkey-male-old
'An old monkey'

c. *ngmàà-dà-kóré lá mà*
monkey-male-old be 1.SG
'I am an old monkey.'

(14) a. *nímí-è áyì bá mùùl-ó tókókyìlé*
eye-PL two NEG peep-IMP baobab.fruit
'You do not use two eyes to look into a baobab.fruit (like a bottle).'

b. *nímí-è áyì*
eye-PL two
'Two eyes'

c. *nímí-è áyì sòbɔ́ lá*
eye-PL two person FOC
'That is a person with two eyes.' = but this can also mean having supernatural
powers.'

(15) a. *báá-ré áyì sóɔ́ngáá bá bɔ̀-rɔ́*
dog-PL two rabbit NEG lose-IMP
'When two dogs chase a rabbit, it hardly escapes.'
ENG: Two heads are better than one: reference to the virtues of teamwork.

b. *báá-ré áyì sóɔ́ngáá á*
dog-PL two rabbit matter FOC
'Two dogs' rabbit'

c. *báá-ré áyì sóɔ́ngáá yélé lá*
dog-PL two rabbit matter FOC
'It is a case of two dogs' rabbit'

(16) a. *wá-bɔ́ng-nɔ́ lá tèɛ̀-rɛ́ kùrí*
NEG-know-AG TOP shoot-IMP tortoise
'It is only a novice hunter who shoots at the tortoise.'

b. *wá-bɔ́ng-nɔ́*
NEG-know-AG
'Novice'

c.   *má wá-bɔ́ng-nɔ́ mɔɔ lá wɛ̀*
NEG-know-AG PART FOC PART
'I see he or she is a novice.'

(17)   a.   *sá á ná   lá pù o ̀ró wò ngó*
stranger TOP greet-IMP deaf
'It is only a stranger who greets the deaf'

b.   *sá á ná   lá pù ò-ró   wò ngó*
stranger TOP greet-IMP deaf
'It is only a stranger who greets the deaf'

c   *sá á ná né ò   wò ngó lá*
stranger and 3.SG deaf   FOC
'It is a case of a stranger and the deaf.'

(18)   a   *bà bá pù g-ró   bà-zò-rá à   ò ní gé só gó*
they NEG praise-IMP   dog-run-AG it face in front
'You do not praise a fast dog in its presence.'

b.   *bà-zò-rá à*
dog-run-AG
'A fast dog'

c.   *à ng lá   pù g-ró bà-zò-rá à   ò ní gè só gó ?*
who FOC praise dog-run-AG 3.SG face in front
'Who praises a fast dog in its presence?'

(19)   a   *zú-ngmá á-rá lá zò-ró   sò ɔ́*
head-cut-AG   TOP fear-IMP   knife
'The executioner fears the knife'

b.   *zú-ngmá á-rá*
Head-cut-AG
'Headcutter'

c.   *má fò   bɔ́ng lá   bɔ́n ná zú-ngmá á-rá náng zò-ró*
PART 2.SG know FOC thing that head-cutter know fear-IMP
'I suppose you know what the executioner fears.'

The above process of shortening proverbs may even lead to new word formation. The following proverb in (20a) can actually serve as input to a lexical operation in which the output is a new word in the language and has a slightly different meaning from the original proverb as in (20b). This shows that lexicalized items may result from (proverb) shortening.

(20) a. *bì-bílé sèé báá*
child-small better than dog
'A Child is better than a pet'

   b. *bìbìl-sèɛ̀-bàá*
child-small-better-than-dog
'A quite grown child'

4. Summary and conclusion

In this paper, we have discussed the phenomenon of linguistic form compression in the Dagaare language of northern Ghana. While many types of word classes and expressions lend themselves to this process, we have discussed just three types here, viz, apocope, compound word compression, and proverb compression. Two main principles have been found to govern the morphosemantic process of linguistic form compression. These are the principle of semantic equivalence and the principle of salience. The former imposes a constraint of semantic equivalence between the inputs and outputs to the process, while the later actually picks out the best candidate from a list of eligible ones that satisfy the principle of semantic equivalence.

Further investigation of the phenomenon of linguistic form compression, both within individual languages and across them, should enable linguists to discover yet other principles – grammatical, pragmatic, and sociolinguistic - that further constrain and explain the process.

5. References

Anttila, A. and Adams Bodomo. 1997. Stress and tone in Dagaare. Manuscript. Boston University.

Bodomo, A. B. 1997. *The Structure of Dagaare*. Stanford Monographs in African Languages. CSLI Publications

Cahill, Mike. 1999. *Aspects of the Morphology and Phonology of Konni*. Ph.D. dissertation, Ohio State University, Ohio, USA

Hyman, Larry. 1993. *Structure Preservation and Postlexical Tonology in Dagbani. Phonetics and Phonology*, vol. 4: Studies in Lexical Phonology, ed. By Sharon Hargus and Ellen Kaisse, 235 - 54. San Diego, California: Academic Press.

Kennedy, J. 1966. *Collected Field Reports on the Phonology of Dagaari*. Collected Language Notes 6. Institute of African Studies, University of Ghana, Accra

Luke, K. K. and O. T. Nancarrow. 1999. Short forms in Hong Kong Cantonese. Paper read at the 7th International Conference on Cantonese and other Yue Dialects, June 26 - 28, 1999, University of Hong Kong.

Rialland, Annie. 1985. Le fini/infini ou l'affirmation/l'interrogation en moba (language voltaique parlee au nord-togo). *Studies in African Linguistics*, Supplement 9, 258 - 61.