Making Athabaskan Dictionaries Usable*

William J. Poser
University of Pennsylvania

Designing a dictionary for an Athabaskan language presents unusual difficulties. Because of the enormous complexity of the verb, it is impossible to list every form of every verb. Because Athabaskan languages combine extensive prefixation with complex stem variation, and because the components that contain the basic meaning of the verb are distributed throughout the form, intercalated with grammatical morphemes, there is no straightforward, easily extracted and manipulated, citation form. Using a fixed member of the paradigm is also problematic because the user must have substantial knowledge of the language to be able to convert other forms to the citation form. As a result, dictionary designers have had two unpleasant choices. One is to use fully inflected forms. These are easy to use, but necessarily far from complete. The other possibility is to produce root-based analytic dictionaries. Such dictionaries may be comprehensive but are almost impossible to use for anyone without considerable meta-knowledge of the language.

The way between the Scylla of incompleteness and the Charybdis of unusability is an on-line dictionary, internally analytic, with a morphological parser as front end. This will allow the user to enter a fully inflected word to be analyzed by the parser. However, difficult problems arise as to how to present the information generated by such a system. Just as finding a word in an analytic dictionary is not trivial, so is making use of the output from one.

* I am grateful to Steven Bird, Catherine Coldwell, Martin Kay, Mark Liberman, Mike Maxwell, and Archie Patrick for discussion of these issues.

1. Introduction

Although dictionaries of some sort exist for many of the Athabaskan languages, in most cases these dictionaries are very far from comprehensive. This is true not only of pocket dictionaries (e.g. Wilson 1989) and dictionaries intended for elementary school classes, which are small by design, but also of dictionaries intended to be as comprehensive as possible, with thousands of entries occupying hundreds of pages, e.g Bray (1998) and Poser (1999, 2000, 2001). The reason for this is that entering words in an Athabaskan dictionary does not work. Because of the nature of Athabaskan verbal morphology, a dictionary with full words as entries cannot be comprehensive. Furthermore, it will fail to reveal information about the relationships among words that is useful to the user, and it will be difficult to include all of the detail that is desirable.

As a result of this problem with verbs, dictionaries have been created for Athabaskan languages of types rarely found for other languages. Since the main difficulty is created by verbs, the problem of creating a full dictionary has sometimes been deferred by publishing a noun dictionary, one that contains only nouns, or sometimes, nouns as well as other non-verbal items, such as adverbs and postpositions. Examples are: Hargus (1986) for Sekani, Kaska Tribal Council (1997) for Kaska, Naish, Story and Davis (1976) for Tlingit and Rice (1977) for Slavey. Some other dictionaries, such as Saxon and Siemens (1996) for Dogrib, while not noun dictionaries in the strict sense, consist overwhelmingingly of nouns, with a rather small number of verbal entries.

2. Athabaskan Morphology

Since the reason that Athabaskan languages present such difficulties is the morphology of the verb, we begin with an explanation of the relevant properties of the verbal morphology. I will illustrate using examples from the Stuart Lake dialect of Carrier, the native language of much of the central interior of British Columbia.

The central problem with Athabaskan verbs is the very large number of forms that they may take on, which virtually precludes listing all of the forms of a verb in the dictionary,

together with the lack of a usable citation form. The exact number of forms of a verb varies from language to language and is difficult to compute, but it is not difficult to see that the number is large.

In (1)-(4) I present a partial paradigm of "to eat". There are nine subject forms, for first, second, and third persons in the singular, dual, and plural.[1] There are four modes: imperfective (I), perfective (P), future (F), and optative (O). Each of these may occur in the affirmative (A) or negative (N). Each verb therefore has a basic paradigm of 72 forms.

With a transitive verb such as "to eat", there is a form used with an overt object, the specified object form (e.g. in "I am eating bread"), and there is another form used when no object is mentioned, the unspecified object form (e.g. in "I am eating"). Yet another distinction is between habitual and non-habitual actions. The forms here labelled "habitual" are used in reference to a typical instance of a habitual act, such as eating lunch. Since the dimensions of habituality and specification of the object are independent, we may combine them freely, leading to a total of 288 verb forms.

Nor do these exhaust the possible forms of the verb "to eat". For example, there are forms used to indicate that the action is performed customarily, and there are forms used to indicate that the object (in this case, what is eaten) is round, is stick-like, or is areal. These two possibilities alone multiply the number of possible forms by eight, for a total of 2304. In addition to the specified object form (with no object marker) and the unspecified object form, there are forms for other objects, such as *sʌʔaɬ* "he is eating me", *yʌʔaɬ* "he is eating it", and *bʌsʔaɬ* "I am eating them", all of which, in principle, may take the shape classifier prefixes. When we combine these with the other categories, we obtain a total of 9,216 forms.[2] Furthermore, from "to eat" we may make a causative "to feed", which takes an indirect object, which may also be marked on the verb. These causative forms therefore number at least 9,216 times 7, or 64,512.. The basic "eat" forms together with the causatives number in total 73,728. These by no means exhaust the possibilities. In short, a Carrier verb may have tens, if not hundreds, of thousands of forms.

---

[1] In this, as in most Carrier verbs, the second person dual and plural are the same and the third person dual and plural are the same.

[2] This number is actually on the low side since we here ignore some of the complexities of the system, such as the existence of reciprocal forms when the subject is plural, with means like "they are eating each other".

(1) To Eat (Unspecified Object)

IA	sg.	du.	pl.
1	ʔʌsʔaɬ	ʔit'aɬ	ʔʌts'aʔaɬ
2	ʔinʔaɬ	ʔahʔaɬ	ʔahʔaɬ
3	ʔaʔaɬ	ʔahaʔaɬ	ʔahaʔaɬ

IN	sg.	du.	pl.
1	ɬeʔzʌsʔal	ɬeʔzit'al	ɬeʔts'ʌsʔal
2	ɬeʔzinʔal	ɬeʔzahʔal	ɬeʔzahʔal
3	ɬeʔʌsʔal	ɬeʔhʌsʔal	ɬeʔhʌsʔal

PA	sg.	du.	pl.
1	ʔisʔal	ʔat'al	ʔʌts'anʔal
2	ʔanʔal	ʔihʔal	ʔihʔal
3	ʔanʔal	ʔʌhanʔal	ʔʌhanʔal

PN	sg.	du.	pl.
1	ɬeʔʌsʔʌɬ	ɬeʔit'ʌɬ	ɬeʔts'iʔʌɬ
2	ɬeʔinʔʌɬ	ɬeʔahʔʌɬ	ɬeʔahʔʌɬ
3	ɬeʔiʔʌɬ	ɬeʔhiʔʌɬ	ɬeʔhiʔʌɬ

FA	sg.	du.	pl.
1	ʔʌtisʔʌɬ	ʔʌtat'ʌɬ	ʔʌztiʔʌɬ
2	ʔʌtanʔʌɬ	ʔʌtihʔʌɬ	ʔʌtihʔʌɬ
3	ʔʌtiʔʌɬ	ʔotiʔʌɬ	ʔotiʔʌɬ

FN	sg.	du.	pl.
1	ɬeʔtʌzisʔʌɬ	ɬeʔtʌzat'ʌɬ	ɬeʔts'ʌtisʔʌɬ
2	ɬeʔtʌzanʔʌɬ	ɬeʔtʌzahʔʌɬ	ɬeʔtʌzahʔʌɬ
3	ɬeʔtisʔʌɬ	ɬeʔhʌtisʔʌɬ	ɬeʔhʌtisʔʌɬ

OA	sg.	du.	pl.
1	ʔusʔaɬ	ʔot'aɬ	ʔʌts'uʔaɬ
2	ʔonʔaɬ	ʔuhʔaɬ	ʔuhʔaɬ
3	ʔuʔaɬ	ʔʌhuʔaɬ	ʔʌhuʔaɬ

ON	sg.	du.	pl.
1	ɬeʔzusʔal	ɬeʔzot'al	ɬeʔts'usʔal
2	ɬeʔzonʔal	ɬeʔzuhʔal	ɬeʔzuhʔal
3	ɬeʔusʔal	ɬeʔhusʔal	ɬeʔhusʔal

(2) To Eat (Specified Object)

IA	sg.	du.	pl.
1	ʌsʔaɬ	it'aɬ	ts'aʔaɬ
2	inʔaɬ	ahʔaɬ	ahʔaɬ
3	aʔaɬ	haʔaɬ	haʔaɬ

IN	sg.	du.	pl.
1	ɬʌzʌsʔal	ɬʌzit'al	ɬts'ʌsʔal
2	ɬʌzinʔal	ɬʌzahʔal	ɬʌzahʔal
3	ɬʌsʔal	ɬʌhʌsʔal	ɬʌhʌsʔal

PA	sg.	du.	pl.
1	ʌsʔal	it'al	ts'anʔal
2	inʔal	ahʔal	ahʔal
3	anʔal	hanʔal	hanʔal

PN	sg.	du.	pl.
1	ɬʌsʔʌɬ	ɬit'ʌɬ	ɬts'iʔʌɬ
2	ɬinʔʌɬ	ɬahʔʌɬ	ɬahʔʌɬ
3	ɬiʔʌɬ	ɬʌhiʔʌɬ	ɬʌhiʔʌɬ

FA	sg.	du.	pl.
1	tisʔʌɬ	tat'ʌɬ	ʌztiʔʌɬ
2	tanʔʌɬ	tihʔʌɬ	tihʔʌɬ
3	tiʔʌɬ	hʌtiʔʌɬ	hʌtiʔʌɬ

FN	sg.	du.	pl.
1	ɬtʌzisʔʌɬ	ɬtʌzat'ʌɬ	ɬts'ʌtisʔʌɬ
2	ɬtʌzanʔʌɬ	ɬtʌzahʔʌɬ	ɬtʌzahʔʌɬ
3	ɬtisʔʌɬ	ɬotisʔʌɬ	ɬotisʔʌɬ

OA	sg.	du.	pl.
1	usʔaɬ	ot'aɬ	ts'uʔaɬ
2	onʔaɬ	uhʔaɬ	uhʔaɬ
3	uʔaɬ	huʔaɬ	huʔaɬ

ON	sg.	du.	pl.
1	ɬʌzusʔal	ɬʌzot'al	ɬts'usʔal
2	ɬʌzonʔal	ɬʌzuhʔal	ɬʌzuhʔal
3	ɬusʔal	ɬʌhusʔal	ɬʌhusʔal

(3) To Eat (Unspecified Object, Habitual)

IA	sg.	du.	pl.
1	naʔʌst'ał	naʔit'ał	naʔts'ʌt'ał
2	naʔint'ał	naʔʌht'ał	naʔʌht'ał
3	naʔʌt'ał	naʔhʌt'ał	naʔhʌt'ał

IN	sg.	du.	pl.
1	nałeʔzʌst'al	nałeʔzit'al	nałeʔts'ʌst'al
2	nałeʔzint'al	nałeʔzaht'al	nałeʔzaht'al
3	nałeʔʌst'al	nałeʔhʌst'al	nałeʔhʌst'al

PA	sg.	du.	pl.
1	naʔist'al	naʔat'al	naʔts'ant'al
2	naʔant'al	naʔiht'al	naʔiht'al
3	naʔant'al	naʔhant'al	naʔhant'al

PN	sg.	du.	pl.
1	nałeʔʌst'ʌł	nałeʔit'ʌł	nałeʔts'it'ʌł
2	nałeʔint'ʌł	nałeʔaht'ʌł	nałeʔaht'ʌł
3	nałeʔit'ʌł	nałeʔhit'ʌł	nałeʔhit'ʌł

FA	sg.	du.	pl.
1	naʔtist'ʌł	naʔtat'ʌł	naʔʌztit'ʌł
2	naʔtant'ʌł	naʔtiht'ʌł	naʔtiht'ʌł
3	naʔtit'ʌł	naʔhʌtit'ʌł	naʔhʌtit'ʌł

FN	sg.	du.	pl.
1	nałeʔtʌzist'ʌł	nałeʔtʌzat'ʌł	nałeʔʌztist'ʌł
2	nałeʔtʌzant'ʌł	nałeʔtʌzaht'ʌł	nałeʔtʌzaht'ʌł
3	nałeʔtist'ʌł	nałeʔhʌtist'ʌł	nałeʔhʌtist'ʌł

OA	sg.	du.	pl.
1	naʔust'ał	naʔot'ał	naʔts'ut'ał
2	naʔont'ał	naʔuht'ał	naʔuht'ał
3	naʔut'ał	naʔhut'ał	naʔhut'ał

ON	sg.	du.	pl.
1	nałeʔzust'al	nałeʔzot'al	nałeʔts'ust'al
2	nałeʔzont'al	nałeʔzuht'al	nałeʔzuht'al
3	nałeʔust'al	nałeʔhust'al	nałeʔhust'al

(4) To Eat (Specified Object, Habitual)

IA	sg.	du.	pl.
1	nast'ał	nait'ał	nats'ʌt'ał
2	naint'ał	naht'ał	naht'ał
3	nat'ał	nahʌt'ał	nahʌt'ał

IN	sg.	du.	pl.
1	nałʌzʌst'al	nałʌzit'al	nałts'ʌst'al
2	nałʌzint'al	nałʌzaht'al	nałʌzaht'al
3	nałʌst'al	nałʌhʌst'al	nałʌhʌst'al

PA	sg.	du.	pl.
1	naist'al	naat'al	nats'ant'al
2	naant'al	naiht'al	naiht'al
3	naant'al	nahant'al	nahant'al

PN	sg.	du.	pl.
1	nałʌst'ʌł	nałit'ʌł	nałts'it'ʌł
2	nałint'ʌł	nałaht'ʌł	nałaht'ʌł
3	nałit'ʌł	nałʌhit'ʌł	nałʌhit'ʌł

FA	sg.	du.	pl.
1	natist'ʌł	natat'ʌł	naztit'ʌł
2	natant'ʌł	natiht'ʌł	natiht'ʌł
3	natit'ʌł	nahʌtit'ʌł	nahʌtit'ʌł

FN	sg.	du.	pl.
1	nałtʌzist'ʌł	nałtʌzat'ʌł	nałʌztist'ʌł
2	nałtʌzant'ʌł	nałtʌzaht'ʌł	nałtʌzaht'ʌł
3	nałtist'ʌł	nałotist'ʌł	nałotist'ʌł

OA	sg.	du.	pl.
1	naust'ał	naot'ał	nats'ut'ał
2	naont'ał	nauht'ał	nauht'ał
3	naut'ał	nahut'ał	nahut'ał

ON	sg.	du.	pl.
1	nałʌzust'al	nałʌzot'al	nałts'ust'al
2	nałʌzont'al	nałʌzuht'al	nałʌzuht'al
3	nałust'al	nałʌhust'al	nałʌhust'al

If these thousands of forms could be subsumed under a single citation form, constructing an Athabaskan dictionary would not be such a problem. The difficulty arises from the conjunction of the large number of forms and the way in which the pieces of the verb are arranged. In English, the singular of nouns and the infinitive of verbs provide useful citation forms, under which other forms may be listed. This is because these citation forms are in most cases easily extracted from the other forms, even by people with little knowledge of English morphology, and because the usual beginning-to-end alphabetic order results in morphologically related forms being, if not adjacent, near each other.

In Athabaskan languages, there is in general no contiguous invariant portion of the verb that can serve as the citation form. The morphology is primarily prefixal, but the existence of extensive stem variation and some suffixation means that the stem is not a good citation form, and that ordering forms form end-to-beginning will not do a good job of keeping related forms close together. What is worse, the phonological material that contributes the basic meaning of the word is not, in general, contiguous. This means that any citation form will necessarily be rather abstract, not easily extracted by an unsophisticated user. Moreover, no simple form of sorting will keep related forms together.[3]

The verb "to eat" illustrated above is in a sense simple, in that the basic meaning of the word is concentrated in the stem, at the end of the word, with the preceding material marking the subject, object, noun class, tense, mode, aspect and so forth. In general, however, Athabaskan languages sandwich such inflectional material between the stem and "derivational" prefixes. Indeed, in many cases the meaning of the combination of the stem and prefix requires a distinct English word. Consider, for example, the form *dahʌnindil* "they entered (walking)". The stem, *dil*, is actually the stem of the verb meaning "for three or more to walk on one pair of limbs". That this form means "to enter" results from the presence of the prefix *da* "through a portal". Together the stem and the prefix indicate that three or more people walked through a portal. If we wanted to say that people had entered a boathouse on board a boat, we would instead say *dahʌninki*, with the same prefix, subject marker, and tense marker, but the stem for going by boat.

What makes matters even worse is that many verb roots are quite abstract, so that a form can only be given an English translation on the basis of the root together with one or more prefixes. We have already seen a minor example of this type in "they entered". Another large set of examples is found in the classificatory verbs. There is a set of verb roots that mean "to handle an object of a certain type in a controlled manner". These are illustrated in (5).

---

[3] It is often said that the fact that Athabaskan languages are prefixal is in and of itself a problem for organizing dictionaries (Hargus 1996;366, Bird, Jeffcoat and Hammond 2001;35). This is not the case. If the basic meaning of the form were contained in an invariant final stem, this stem would not be difficult to extract, and sorting from end-to-beginning, as in "reverse" dictionaries, would suffice to keep related forms close together.

(5) 'he will give me' for Different Types of Object

non-plural default object (chair)	sʏatiʔał
non-plural *n*-class object (ball)	sʏantaʔał
non-plural *d*-class object (name)	sʏadʌtaʔał
non-plural x^{w}-class object (house)	sʏautaʔał
plural default objects (chairs)	sʏatilił
plural *n*-class objects (balls)	sʏantalił
plural *d*-class objects (names)	sʏadʌtalił
plural x^{w}-class objects (houses)	sʏautalił
uncountable generic objects (sugar)	sʏatidzih
uncountable *n*-class objects (berries)	sʏantadzih
uncountable *d*-class objects (toothpicks)	sʏadʌtadzih
long rigid object (canoe)	sʏatiteł
long rigid *d*-class object (stick)	sʏadʌtateł
body (dog)	sʏatiłteł
contents of open container (cup of tea)	sʏatikał
2-d flexible object (shirt)	sʏatiłčʌs̲
mushy stuff (mud)	sʏatitloh
liquid (water)	sʏatiłdzo
hay-like (hay)	sʏadʌtałdzo
fluffy stuff (down)	sʏantałdo

These handling stems combine with prefixes to yield verbs describing particular types of handling. In (6), for example, we illustrate a few of the handling verbs for two-dimensional flexible objects, such as shirts.

(6) Different Types of Handling of a Single Type of Object

behanaitiłčʌs̲	he is going to take it out
didʌtałčʌs̲	he is going to hold it up
dʌʏaitałčʌs̲	he is going to hang it up
k'itałčʌs̲	he is going to put it on (the table)
k'ʌnaitałčʌs̲	he is going to put it back on (the table)
k'ʌnaitiłčʌs̲	he is going to take it off (the table)
sanaitiłčʌs̲	he is going to bring it back
yʌʏatiłčʌs̲	he is going to give it to her
yʌʏʌtiłčʌz̲	he is going to lend it to her
nʌtiłčʌz̲	he is going to carry it around
ʔatiłčʌs̲	he is going to bury it
tatiłčʌs̲	he is going to submerge it
natiłčʌs̲	he is going to put it on the ground
yaiyʌtiłčʌs̲	he is going to bring it ashore

To take yet one other example, the same verb root, meaning something like "to make an abrupt motion", underlies "to pick (flowers)", "to pluck (a chicken)", "to weed (the garden)", "to snap (a stick) into two pieces)" and "to scrub (on an old-fashioned washboard)".[4]

A final difficulty is that many verbs consist of a discontinuous combination of a stem and one or more prefixes. In (7) we have the future negative paradigm of the verb "to speak".

(7) To Speak [FN]

	singular	dual	plural
1	yałtʌzisdʌk	yałtʌzaldʌk	yałts'ʌsdʌk
2	yałtʌzałdʌk	yałtʌziłdʌk	yałtʌziłdʌk
3	yałtisdʌk	yałotisdʌk	yałotisdʌk

The subject markers (1s /s/, 2s /in/, 3s zero, 1d /id/, 1p /ts'/, 2dp /h/ and 3dp /hV/), the negative markers (/ł/ and /s/), and the future tense markers (inceptive /t/ and aspect vowel /i/, which fuses with the initial /i/ of the 2s and 1d subject markers to yield /a/) all come in between the prefix *ya* and the stem *dʌk*, just as they do in "to enter". Here, however, no meaning can be assigned to the prefix *ya*. Part of one's knowledge of Carrier is that "to speak" calls for the stem *dʌk*, the immediately preceding prefix *ł*, and the prefix *ya* way out near the left edge.

Another example arises with the handling verbs already discussed. The classifier prefix /d/ mentioned in (5) is generally used with long, thin, stick-like things. However, it is also used, in combination with the default handling verbs, in reference to rocks. One simply has to know that the verb forms appropriate for handling rocks consist of the default verb bases, as one might expect, together with the /d/ classifier prefix, which with all other verbs would be inappropriate. Such discontinuous verb themes are common.

These morphological characteristics of Carrier and other Athabaskan languages make it particularly difficult to create dictionaries for them. The vast number of verb forms means that it is impossible to record all of them, and if one could, it would not possible to print them in a volume of reasonable size. The fact that the pieces of the verb that convey the basic, non-grammatical meaning may be discontinuous, separated by grammatical morphemes, means that there is no easily extracted, eontiguous citation form, comparable to the English infinitive, under which the various forms may be listed.

3. Traditional Approaches

One approach to this problem has been to list, for each verb, one agreed upon form. This is the approach of the major dictionary of Navajo (Young and Morgan 1987). However, this requires the user to be able to analyze the verb form he or she has heard or read and to convert it to the citation form. This is a non-trivial task even for fluent native speakers of the language; it is difficult or impossible for language learners.

[4] I suspect that this is also the root underlying"to hiccup".

The other major approach is not to list fully inflected words at all, but to list individual morphemes. In order to use such a dictionary, the user must be able to analyze the word and in particular to identify the root.

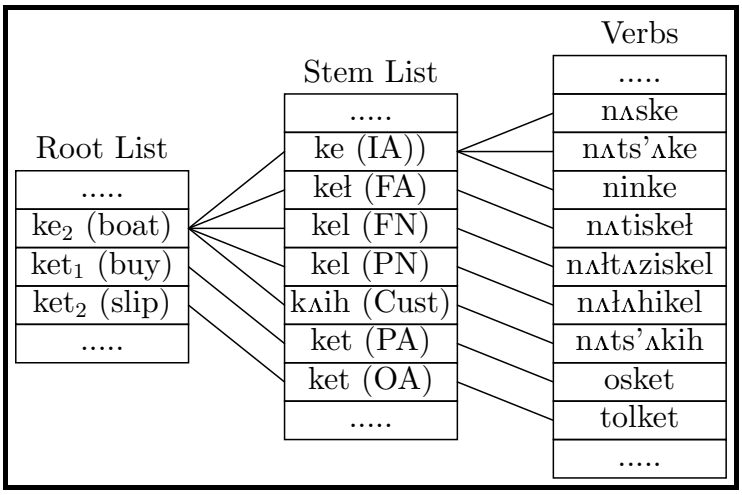
Identifying the root is not as simple as it might seem because of the complex relationship between roots and stems in Athabaskan languages. Athabaskan verbs consist of a stem, approximately the last syllable, and a set of prefixes. In general, the stem carries the main meaning of the verb. In (8) we have forms representing the tense/mode/negation paradigm of the verb "to go around in a boat". For each form, the last syllable, which is the stem, is shown separated from the remainder, which consists of prefixes. In this case the prefixes convery information about the subject, tense/mode, and negation, except for the initial /n/, which denotes motion in a loop. We see that the stem varies considerably in form with tense, mode, and negation.

(8) Stems of "I go around in a boat"

TMA/Neg	Prefixes	Stem
Imperfective Affirmative	nʌs	ke
Imperfective Negative	nʌɬʌzʌs	koh
Perfective Affirmative	nʌsʌs	ki
Perfective Negative	nʌɬʌs	kel
Future Affirmative	nʌtis	keɬ
Future Negative	nʌɬtʌzis	kel
Optative Affirmative	nos	keʔ
Optative Negative	nʌɬʌzus	keʔ

These various stems are all associated with an abstract root, in this case *ke*, meaning "go by boat", from which they are considered to be derived. Although there is a pattern to the changes in stems, it is complex if not irregular, and so someone learning the language must to a considerable extent simply memorize the stem set for each verb. The relationship among individual verb forms, stems, and roots is illustrated in (9).

(9) Relationships of Verbs, Stems, and Roots in a Carrier Dictionary



To look up a word in an analytic dictionary, the user first has to identify the stem and look up what stems of what roots it might be. The user then looks up each possible root, under which information is presented about possible derivations. This information together with the user's knowledge of the grammar in theory enables him or her to work out the meaning of the whole form from those of its parts.

Such analytic lexica have been published for Navajo (Young, Morgan and Midgette 1992), Ahtna (Kari 1990), and Koyukon (Jetté and Jones 2000). They have the virtue, of being, in principle, comprehensive. Furthermore, they allow detailed information to be provided without duplication. For example, the detailed meaning of a verb root can be explained only once, in the entry for that root, rather than in each of many entries for forms derived from that root.

The problem with this approach is that it requires a considerable amount of grammatical knowledge on the part of the user, together with an understanding of a fairly elaborate process for analyzing forms, looking up their components, and constructing the meaning of the form from its components. As a result, analytic dictionaries are fine for linguists knowledgable about a language, but experience with the existing analytic dictionaries shows that for most people, including both language learners and native speakers without considerable metalinguistic knowledge, they are very difficult to use.

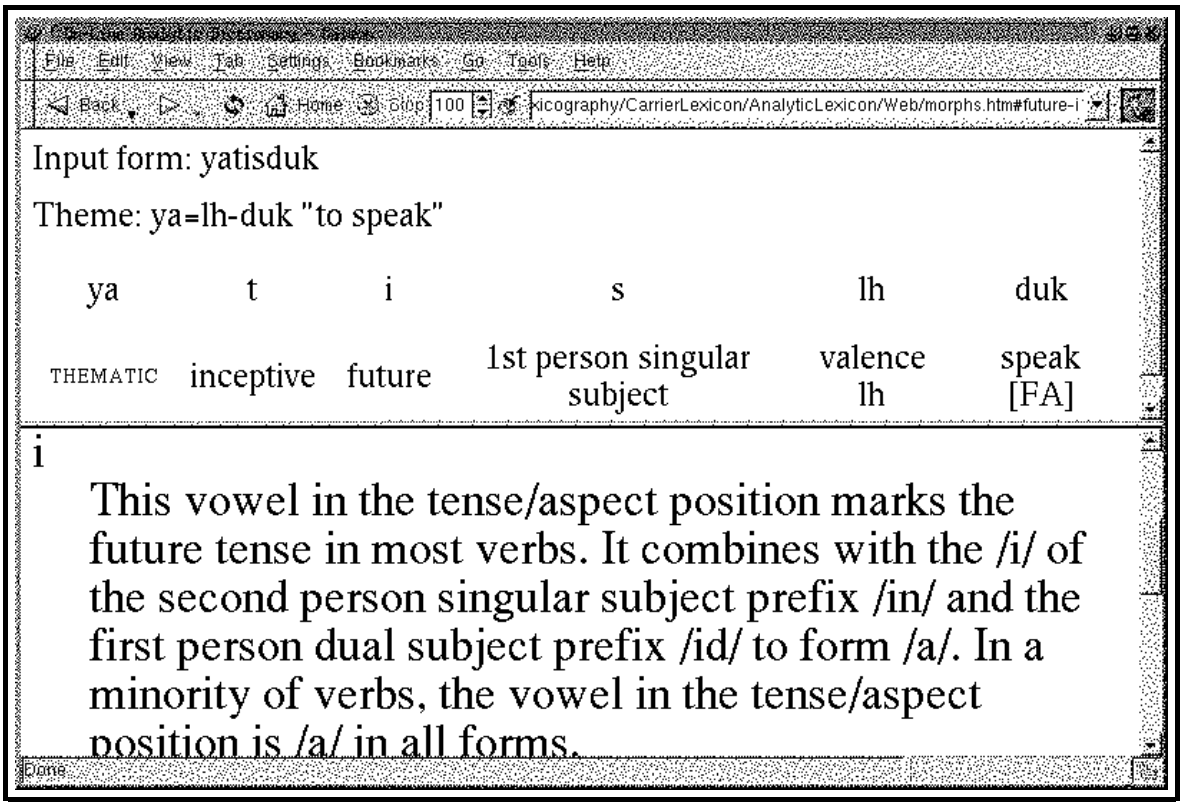
4. On-Line Analytic Dictionaries

The solution that presents itself is to create a dictionary that is at heart an analytic lexicon, but to eliminate the need for the user to have an extensive knowledge of the morphology and to be able to analyze words. This can in theory be done by using a morphological parser as a front end to an on-line analytic lexicon. This means that the user would enter a complete word by typing it into a computer containing the same information as a printed analytic dictionary. A computer program would analyze the word, identifying the root and the affixes, and look them up in the dictionary.

A lexicon with a morphological parser as a front end is not a complete solution. For a language specialist, it may be sufficient to know, as the parser will report, that *nʌtiskeł* consists of the prefixes *n* "in a loop" *t* "inceptive" *i* "future" *s* "first person singular subject" and the stem *keł* "go by boat" (Future Affirmative), but to understand that this means "I am going to go around in a boat" requires an understanding of the grammar together with the ability to combine the meanings of the various pieces.

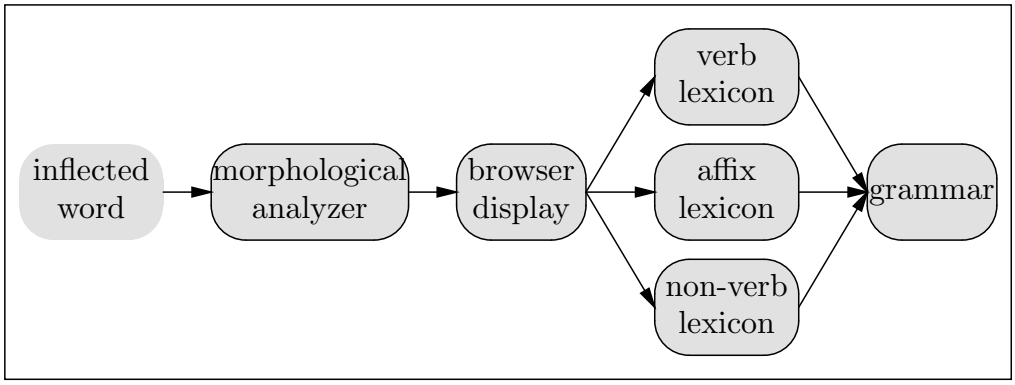
A partial solution to this problem is for the output of the dictionary program to consist of a piece of hypertext, giving the analysis of the word, with each morpheme a link to the lexical entry for that morpheme. This approach is illustrated in (10), which depicts a browser display of a mock-up of the output of an on-line analytic dictionary. In (10) the user has selected the link associated with the future morpheme /i/ in the parse in the upper frame, resulting in the display of the information in the lower frame.

(10) A Browser Display after a Selection



These lexical links might in turn lead to the appropriate sections of the reference grammar. The overall structure of such a system is shown in (11).

(11) The Structure of an On-Line Analytic Dictionary



This approach, however, does not completely solve the problem. It relieves the user of the need to analyze the input form, but it does not really eliminate the need for the user to construct the meaning of the form as a whole from information about the parts. A still more sophisticated system might attempt to synthesize the meaning of the form for the user, generating an English word or phrase as output, in place of, or in addition to, the analytic output described above.

Although this is easy to describe, it is problematic in practice. One reason is that experience with text synthesis has shown that, except for very restricted domains (such

as generating reports from stock market information), generating text from a semantic representation is very difficult to do.

The other reason is that English words, and even English phrases, simply do not correspond one-to-one to Athabaskan words and phrases. It may be possible to convey the same information in one language as in the other, in some cases by means of lengthy explanations, but it is in general not possible to translate words or even phrases between Athabaskan languages and English. Because the languages are so different, even with a good text generation component, it will always be necessary to some extent to tell the user what the components of the words are and to use technical vocabulary to explain their usage and meaning.

Let us take one simple example. Like the other Athabaskan languages, Carrier has a form of the verb that linguists call the "optative". This form is so-called because, among other things, it may be used to express the speaker's wish. Optative forms with first person dual or plural subjects can be used to say "let's do such-and-such". *naʔtsʼutʼaɬ*, for example, can mean "let's eat". Second person optative forms can be used as a more polite form of imperative. Third person optative forms can be equivalent to English expressions such as "let him/her do such-and-such". *tuyaʔ* can mean "let him go". Even here we see that the translation will be complex, because English uses different expressions in different persons for what is arguably the same idea.

In fact, the situation is much worse, because this is only one of many uses of the optative. Negative commands are made with the affirmative form of the optative followed by a negative particle. The optative affirmative followed by the complementizer *xʷʌčʼa* constitutes a "lest" clause, as in (12):

(12) Nγu dułjʌt xʷʌčʼa tʌnaingʌs.
your-teeth decay-OA lest you-wash
Brush your teeth lest they decay.

The optative is also used in the lower clause in the "tell to" construction:

(13) Dʌyeʔ hutʼen wxʷe xʷʌsanaujaʔ yʌɬni.
his-son it-is-light while he-returns-OA he-told-him
He told his son to return while it is still light.

There is arguably something that unifies these various uses of the optative, namely that they are all *irrealis*, that is, that they have to do with situations whose occurence is uncertain, but there is no single English construction that captures all of them. The Carrier optative simply cannot be translated into English except in a sufficiently well-defined context. What a dictionary can tell someone about an optative form is what the meaning of the verb is and that it is an optative form. Any further explanation can only consist of information about what the optative form is good for.

It seems, therefore, that in the short term the best approach to generating output from an on-line analytic lexicon is present the parse in as helpful a way as possible and to allow the user to move easily from the analysis of the word to information about its components. Some of this information will be located in the dictionary; other information will be located in the grammar.

References

Bird, Sonya, Melody Jeffcoat, and Michael Hammond (2001) "Electronic dictionaries for languages of the Southwest," in Steven Bird, Peter Buneman, and Mark Liberman (eds.) *Proceedings of the IRCS Workshop on Linguistic Databases* Philadelphia: Institute for Research in Cognitive Science, University of Pennsylvania. pp. 32-37. `http://www.ldc.upenn.edu/annotation/database`.

Bray, Dorothy (ed.) (1998) *Western Apache-English Dictionary: A Community-Generated Bilingual Dictionary.* Tempe, Arizona: Bilingual Press/Editorial Bilingüe.

Hargus, Sharon (1986) *Sekani Topical Noun Dictionary.* Ms., Department of Linguistics, University of Washington.

Hargus, Sharon (1996) Review of Young, Morgan, and Midgette (1992) *Analytical Lexicon of Navajo*. *Anthropological Linguistics* **38**.2.366-370.

Jetté, Jules & Eliza Jones (2000) *Koyukon Athabaskan Dictionary.* Fairbanks: Alaska Native Language Center.

Kari, James (1990) *Ahtna Dictionary.* Fairbanks: University of Alaska Press.

Kaska Tribal Council (1997) *Guzą́gi K'ų́gé' – Our Language Book: Nouns. Kaska, Mountain Slavey and Sekani.* Whitehorse: Kaska Tribal Council. Two volumes.

Naish, Constance M., Gillian L. Story, and Henry Davis (1976) *Tlingit Noun Dictionary.* Fairbanks: Sheldon Jackson College.

Poser, William J. (1999) *Nak'albun/Dzinghubun Whut'enne Bughuni* (Stuart/Trembleur Lake Carrier Lexicon). Vanderhoof, BC: Yinka Dene Language Institute. Second edition.

Poser, William J. (2000) *S̲aik'uz̲ Whut'en Hubughunek* (Stoney Creek Carrier Lexicon). Vanderhoof, BC: S̲aik'uz̲ First Nation. Fourth edition.

Poser, William J. (2001) *Lheidli T'enneh Hubughunek* (Fort George Carrier Lexicon). Prince George, BC: Lheidli T'enneh. Third edition.

Rice, Keren (1977) *Hare Noun Dictionary.* Ottawa: Northern Social Research Division, Department of Indian and Northern Affairs.

Saxon, Leslie and Mary Siemens (eds.) (1996) *Tl̨ícho Yatíí Enihtl'èe: A Dogrib Dictionary.* Rae-Edzo, Northwest Territories: Dogrib Divisional Board of Education.

Wilson, Garth A. (1989) *Conversational Navajo Dictionary: English to Navajo.* Blanding, Utah: Conversational Navajo Publications.

Young, Robert W. and William Morgan, Sr. (1987) *The Navajo Language: a Grammar and Colloquial Dictionary.* Revised edition. Albuquerque: University of New Mexico Press.

Young, Robert W. Morgan, William Sr. and Sally Midgette (1992) *Analytical Lexicon of Navajo.* Albuquerque: University of New Mexico Press.