THREE APPROACHES TO DESCRIBING CONVERSION[1]

Arto Anttila
Research Unit for Computational Linguistics
The SIMPR Project
University of Helsinki

0. Introduction

This paper will be concerned with a grammatical phenomenon known as conversion. Although we shall be exemplifying from English throughout, our point is of a general nature. It is often the case that persistent residual problems in descriptive work are symptomatic of more fundamental theoretical ones.

The subsequent discussion is in two parts. First, some basic facts and a definition will be presented. Second, three strategies for describing conversion will be discussed and evaluated.

1. Preliminaries

In a standard reference grammar of English conversion is regarded as a "derivational process whereby an item is adapted or converted to a new word class without the addition of an affix" (Quirk et al. 1985:1558). Typical examples include the following:

(1) I love you.
(2) Give my love to Rose.
(3) This bottle contains beer.
(4) I am going to bottle some beer today.

However, there appears to be some hesitation as to exactly where in the grammar this phenomenon ought to be described, and the proliferation of terminology ("functional conversion", "functional shift", "zero derivation") is, in this

respect, at least suggestive (cf. Pennanen 1971).

All grammarians agree on one point: conversion is very productive unlike so much of derivational morphology. Quirk et al. (ibid.) go on to remark that "in colloquial usage (firmly regarded as nonstandard in BrE, less so in AmE), recourse to conversion is especially common". Surprisingly enough, conversion is usually treated in appendices; the authoritative work cited above is no exception. It also seems to be the central vocabulary of English, in terms of frequencies, that usually get converted. This is reflected in the statistics quoted by DeRose (1988:31): only 11.5% of English word form types are categorially ambiguous, but the figure for tokens is closer to 40%.[2]

The directionality of conversion warrants a remark. People indeed seem to have intuitions about which item is "basic", which "derived".[3] Quirk et al. (ibid.) suggest two types of diagnostic test (among others). Firstly, the semantic dependence of one item upon another: the verb net can be paraphrased 'catch in a net', but no comparable paraphrase can be given for the noun. Ergo, the noun is basic. Secondly, the noun release is parallel to overtly deverbal nouns as regards semantic restrictions: one may say "His release was sudden/on Thursday" just as one may say "His discovery was sudden/on Tuesday". Ergo, the verb is basic. In this way we arrive at the two traditional classes, denominals (a net -> to net) and deverbals (to release -> a release). Paraphrase and semantic parallelism are difficult tools to work with, but the correctness of the observation is beyond dispute.

We shall now attempt a reasonably neutral definition of conversion:[4]

(5) Definition: Conversion

    Conversion is a relation between two separate lexical
    entries with different parts-of-speech, identical phonological representations and related meanings.

This definition has two important characteristics. First, the type of entity which may enter into a conversion relation is defined as lexical entry. The definition thus excludes syncretism (Huddleston 1984:77), if that is conceived of as a relation between two syntagms sharing the information of two or more lexical entries. Such syntagms are sometimes called morphosyntactic words (Lyons 1977:377):

(6)   <cut:V> + <INFINITIVE>
(7)   <cut:V> + <IMPERATIVE>
(8)   <cut:V> + <PAST>
(9)   <cut:V>
(10) <cut:N>

Thus, only the pair (9)-(10) would be an instance of conversion. The notions 'lexical entry' and 'morphosyntactic word' are, of course, highly description-dependent.[5]

The second major point concerning our definition is that no stand is taken on the implementation issue; conversion is defined simply as a relation. In particular, no kind of derivational relationship between lexical entries is presupposed (cf. section 2.2).

2. Descriptive proposals

In what follows, we shall briefly discuss three different types of solution for describing conversion. It might be of some exegetic interest to reconstruct the views of a few traditional grammarians in our terms. That would, however, take us too far afield and might also be unfair (but see Pennanen 1971).

2.1 A lexical solution: listing

In this type of description (see Lieber 1981), one simply regards all conversion mates as separate lexical entries:

(11) <drink:V>
(12) <drink:N>

There are good arguments for this solution. Typically, they
hinge on the fact that even though one of the conversion
mates is in some sense derived from the other, there are
enough idiosyncracies to seriously weaken any general rule-
based approach (cf. Chomsky 1970). We give two types of
argument:

(i) Semantic idiosyncracy

The nominal and verbal readings of bottle denote something
very prototypically thing-like and process-like, respective-
ly ('a bottle', 'to put in a bottle'). Crucially, the verb
incorporates the meaning 'to put in', and not, for example
'to empty'.   There is an additional, unpredictable meaning
element attaching to the denominal verb, which simply has to
be encoded in its lexical entry (cf. <milk:V> meaning 'to
draw milk out of', not, for example, 'to drink milk'). This
kind of idiosyncracy strongly suggests positing distinct
entries.

(ii) Syntactic idiosyncracy

There are well-known syntactic arguments for postulating
separate entries for derivatives, and the same arguments
would seem to apply to conversion mates as well. The one
based on idiosyncratic prepositions appears particularly
convincing (Horrocks 1987:57). Take the pairs

(13) <answer:V> <answer:N>
(14) <attack:V> <attack:N>

and the sentences

(15) John answered the question
(16) John's answer to the question...

(17) The speaker attacked the proposal.
(18) The speaker's attack <u>against</u> the proposal...

If the preposition were always the same it could be accounted for by a general convention; now, as it seems to vary idiosyncratically from pair to pair the natural way out is to place it as a feature on the noun entry, which implies a lexical split.

Evaluation

(i) Methodology

This solution is, of course, always at hand. Technically, there is nothing to keep one from postulating new lexical entries whenever one feels so inclined. However, for exactly this reason, this seems no solution at all, but rather a sign of resignation. It also appears unnatural in that it leads to a proliferation of homonymy (cf. Lyons 1977:567), or, in parsing terms, lexical category ambiguity.

(ii) Extensive lexical redundancy

Another problem for listing is that there will be numerous pairs of entries which are, both phonologically and semantically, clearly interconnected, but this fact shows up nowhere in our description, which is a definite flaw. The standard cure is some type of a redundancy rule (see Jackendoff 1975) which will serve to pick out the generalization that two distinct entries are related. Redundancy rules have been criticized notably by Hudson (1984:65-72). As he points out, such a family of rules for characterizing the notion 'distinct but related lexical entries' will prove very heterogeneous. At worst, one could envisage a rule type connecting only two individual entries, like <male:A> and <female:A>. Rather, he suggests, there is something wrong with the basic assumption that the lexicon is "an unordered

set of lexical formatives" (Chomsky 1965:84), and he goes on
to argue for a diffuse lexicon with no clear-cut entries.

One function of redundancy rules is to account for the
relatedness of sense between distinct entries. In fact,
there is another device for a similar purpose, namely poly-
semy, i.e. the possibility of incorporating more than one
sense in one and the same lexical entry. It must be asked
whether it is desirable to have both types of device in our
grammar. One might argue that difference of part-of-speech
is a sufficient reason for assuming two entries, since, to
be sure, a single entry labelled N and V simultaneously
would not be polysemous but simply self-contradictory. It
must be remembered, however, that parts-of-speech (as all
lexical classes) are theoretical constructs in need of inde-
pendent justification and their status may well be reasses-
sed (cf. section 2.3).

(iii) Conversion of inflected words

Unfortunately, there is a further problem with cases like

(19)   The oldest were left alone.
(20)   The killed were all young men.

Clearly, one would seem to be obliged to list pairs such as

(21) <oldest:A>   ("The oldest cats swim.")
(22) <oldest:N>   (19)
(23) <killed:V>   ("They were killed.")
(24) <killed:N>   (20)

and, indeed, a vast number of productively inflected items
as lexical entries. Given that conversion is at least semi-
productive, this would lead to a multiplication of lexical
entries beyond any reasonable limit, not to mention loss of
insight.

One way out would be claiming that the NPs above are elliptical, e.g. in (19) <u>the oldest (people)</u>. We will not explore this possibility here. It suffices to point out that ellipses tend to be pragmatically, or at least non-syntacti- cally, conditioned, and in order to validate the approach we should have a sophisticated theory of pragmatics or text linguistics which would make explicit predictions about the occurrence of ellipses. This we do not have for the moment.

2.2 A morphological solution: derivation

This group of solutions has a number of distinguished advo- cates among descriptive grammarians (see for example Jesper- sen 1942:85, Marchand 1969:359-89, Quirk et al. 1985:1558- 67). In most works of reference conversion is placed under derivation. Although the important theoretical discussion concerning the implications of this alternative was carried out within the taxonomic structuralist framework some forty years ago, the problem is very much still with us (for a critical compendium of views, see Haas 1957).

The basic idea is to make conversion line up with derivation proper by postulating zeroes; the standard ar- guments are crucially based on <u>analogy</u>. According to this view, the lexicon itself contains only one of the conversion mates. If <bottle:N> is in the lexicon, the verb gets de- rived by the affixation of a denominal verbalizing morpheme, which, however, is realized as a zero allomorph[6]

(25) <bottle:N> + <..:V:SUBCAT N__>
     /bottle/ + /0/

To justify this, we can point to the existence of a phonolo- gically non-empty (morphologically conditioned) allomorph /ize/, as in <u>alcoholize</u>. Sentences like (19) and (20) would seem to constitute no problem, given that the nominalizing morpheme may attach to inflected words as well (but see the criticism below).

Evaluation

(i) Methodology

Since Pāṇini, zeroes have played a role in linguistic description. However, conjuring up such elements to account for thorny data may be methodologically suspect. If no constraints are placed on zeroes, there is no principled reason for us to stop here. For example, syncretism and unmarked number can be described by assuming entries like

(26) <0:IMPERATIVE:SUBCAT V__>
(27) <0:SINGULAR:SUBCAT N__>

which constitute a special problem, as they are <u>never</u> realized phonologically in any environment. An entity like those in (26)-(27) is sometimes called <u>morphemic</u> zero, as against allomorphic (cf. Nida 1948:46, Gleason 1969:76). Disallowing morphemic zeroes (Bloch 1947, sect. 2.3) will not suffice; as is pointed out by Haas (1957:38), by analogy we can arrive at descriptions like (29):

(28) <lion:N> + <..:FEMALE:SUBCAT N__>
      /lion/   + /ess/
(29) <boy:N> + <..:FEMALE:SUBCAT N__>
      /girl/ + /0/

where /girl/ is a morphologically conditioned allomorph of <boy:N> and the triggering morpheme <..FEMALE..> is realized as a zero allomorph, a somewhat counterintuitive conclusion.
    Like lexical listing, zero derivation is open to the problem of homonymy. What is even worse, the zero allomorph itself (i.e. nothing) will be homonymous many times over, as <..N:SUBCAT V__>, <..V:SUBCAT N__>, <..N:SUBCAT A__> etc. will all have zero as one allomorph. Postulating several different zero allomorphs would avoid homonymy, but only technically so.

(ii) Zeroes are recognizable only redundantly

It has been pointed out (Haas 1957:40) that in order for a
zero allomorph to be of some value, one should be able to
conceive of a situation where its presence could contrast
with its absence. But whenever we are able to "recognize" a
zero, it is by virtue of its syntagmatic context, the zero
itself being "present" only by way of side-effect. As its
distribution is fully conditioned by its environment, one
can rightfully question its status as an independent allo-
morph.

(iii) The order of affixes

There would seem to be a further complication concerning the
order of affixes. The order

(30) ROOT < DERIVATIONAL AFFIXES < INFLECTIONAL AFFIXES

is usually regarded as the norm at least for English and
several other suffix languages (Bloomfield 1933:222); in
prefix languages (e.g. Bantu) the same pattern is mirrored.
Consequently, it has been proposed as a universal that
inflected words cannot be further derived, i.e. inflectional
affixes are not allowed between a stem and a derivational
affix (Greenberg 1963:73, Universal 28).

    But now, zero-deriving oldest as we might be tempted to
do in the case of (19) above seems problematic. If we con-
ceive of affixation in terms of linked minilexica (Kosken-
niemi 1983:27-9), we can start out with <old:A>, enter in-
flection (<est:A>), but further accessing derivation (<0:N>)
would violate the proposed universal order. On the other
hand, first entering derivation (<old:A> + <0:N>) is no
better, since we would certainly not want to allow nouns to
have a free passage to superlative inflection. The problem
is, once again, the status of inflected items which seem to

be eligible for zero-derivation. If oldest cannot be an adjective for syntactic reasons, we must probably conclude that it is syntactically a noun but morphologically an adjective. This is not very satisfactory. The ellipsis solution suggested above would, of course, be immune to this criticism; oldest would simply be an adjective.

2.3 A syntactic solution: recategorization

This type of solution is to our knowledge rarely discussed seriously. It is, however, a possibility. Assume that we have in our lexicon the following entries:

(31) <fly:V>
(32) <ing:V:SUBCAT V__>

but no entries like

(33) <ing:N:SUBCAT V__>
(34) <ing:A:SUBCAT V__>

The problem we immediately face is manifest in (36)-(37):

(35) We are flying.
(36) Flying is dangerous.
(37) Tom saw a flying plane.

In other words, flying shares the distribution of nouns and adjectives as well, an example of triple conversion. Of course, we might resort to lexical listing and posit entries like <flying:V>, <flying:N> and <flying:A>, but this would multiply the size of the lexicon. Alternatively, we could add (33) and (34) to the lexicon. This would seem to have exactly the desired effect, the ambiguity of flying being localized in the suffix. Unfortunately, this would make /ing/ an allomorph of three distinct suffixes. Technically, this shortcoming could be remedied by morphemic zeroes like

(38) <0:N:SUBCAT ING__>

but for reasons given in 2.2 this tack might not be easy to defend. However, there are still other alternatives.

Sloppy syntax

Now, we may start to have doubts that we are just trying to preserve our preconceived idea of phrase structure intact. We have taken for granted that there ought to be rules like

(39) NP --> DET A N

but no rules like

(40) NP --> V   (cf. 36)
(41) NP --> DET V N   (cf. 37)

and, to be sure, rules like (40) and (41) would generate, among others, NPs like those in (36)-(37). The problem is, of course, that both rules overgenerate wildly, not to mention that (40) would be declared illegal by the X'-convention.

There is something basically wrong with this approach. We seem to have forgotten that the raison d'être of word classes (like our V, DET etc.) is to act as pointers to the rest of the grammar by indicating the syntactic equivalence of certain words. Consequently, these classes should be so established as to facilitate stating syntactic generaliza-tions. Instead, we have simply taken a set of pre-establis-hed part-of-speech labels and labelled words more or less intuitively, just to find ourselves patching up the results either by lexical listing, zero suffixes or syntactic rules. It is arguable that formulating discovery procedures for parts-of-speech (Harris 1946) was after all no idle pastime. The traditional ready-made word classes often seem to cloud the facts and lead to complicated syntactic statements.

Recategorization

Having rejected listing, zero morphemes and sloppy syntax, we may now conclude that the crux is rather the word classes themselves. Establishing word classes is not "a terminological preamble" which can be completed "before going on to the 'meaty' part of a grammar" (Crystal 1967:25). Neither can multiple classification be dismissed as a residual problem to be handled as conversion (homonymy/lexical ambiguity).

Reinterpreting conversion as a syntactic (rather than lexical or morphological) phenomenon turns on the fundamental insight that there is no pre-theoretically given boundary between syntax and lexicon. To illustrate this, let us consider two simple examples. First, take the sentence

(42)   We are eating apples.

which manifests a global ambiguity. Now, there are at least two possible ways to describe this ambiguity:

(a) The ambiguity is <u>lexical</u> and due to the word-form <u>eating</u> which is ambiguous as to N, V or A. In our syntax, we would allow for two different strings of parts-of-speech (by whatever means):

(43) PRON AUX A N   (44) PRON AUX V N

which represent the disambiguated readings. Given the sentence (42) and (43) as its syntactic specification, the interpretation is uniquely identifiable ('We are apples which eat'). Note that there is no need to further recognize different phrase structures like

(45) (PRON) (AUX) (A N)
(46) (PRON) (AUX V) (N)

(b) We posit INGFORM as the part-of-speech of eating (cf.
Huddleston 1984:83). Consequently, we have a unique string

(47) PRON AUX INGFORM N

only do we have to distinguish the readings in terms of
phrase structure, i.e. the ambiguity is syntactic:

(48) (PRON) (AUX) (INGFORM N)
(49) (PRON) (AUX INGFORM) (N)

As for another example (see Lyons 1977:400-3), consider the
following:

(50) He hit the man with a stick.

which is a textbook example of syntactic ambiguity:

(51) (He) (hit (the man) (with a stick)).
(52) (He) (hit (the man (with a stick))).

However, with a little imagination this can be seen as an
instance of lexical ambiguity. Instead of having a single
entry <with:PREP> we might opt for homonymy:

(53) <with:PREP INSTRUMENTAL>
(54) <with:PREP COMITATIVE>

and the bracketing is superfluous, just as in (43)-(44).

Evaluation

Compared to its alternatives discussed in sections 2.1 and
2.2, the syntactic solution has certain advantages. For
example, only having entries like

(55) <ing:INGFORM>
(56) <with:PREP>

would not give rise to the problem of homonymy (or, in
parsing terms, lexical category ambiguity), nor would there
be any need for zero allomorphs. If concreteness is one of
our methodological desiderata, the syntactic solution ranks
high: <eating:INGFORM> (or <ing:INGFORM:SUBCAT V__> for that
matter) is seen as one coherent sign (see Nyman 1989:35-9),
just as its unique phonological form would suggest, and not
as a conglomeration of three incompatible morphosyntactic
words which just happen to fall together phonologically. By
syntactic recategorization a nice one-to-one mapping between
form and meaning can be preserved.
    Unfortunately, there are problems as well:

(i) The number of primitives

The price for the resulting neat lexicon and morphology must
be paid in syntax. Introducing new classes like INGFORM
will be done at the cost of generality. If the old classes
still remain in the grammar the number of primitives will
grow larger, not to mention that the category INGFORM is
language-specific (Huddleston 1984:84). This, however, is
not a damaging criticism. Obviously, we cannot decide in
advance what categories in our description language must be
universal and what not.

(ii) Redundancy in syntactic rules

Class labels being pointers to the rest of the grammar, new
labels necessitate a more complicated syntax. Thinking in
terms of PS-rules, we would probably find ourselves making
statements like

(57) NP --> DET A N
(58) NP --> DET INGFORM N

and, were we tempted to continue with recategorization, we
certainly would not want to see <red:A> and <eating:INGFORM>
as members of the same syntactic class. Clearly, having to
make syntactic statements like (57)-(58) shows that further
refinement is due. At least, it must be shown precisely in
what way the classes A and INGFORM (given that we want such
classes) are related.

3. Conclusion

It has not been a part of our purpose in this paper to argue
for any particular solution. Rather, we have simply presen-
ted a number of arguments which to us seem pertinent and
useful as instruments for any description of conversion.
However, it may have become obvious to the reader that our
sympathies lie with the syntactic approach. In our view, the
potential of this alternative is underexplored, and it seems
worthwhile giving it serious thought.

Notes

[1] I wish to thank Martti Nyman for useful comments. Of
course, the usual disclaimers hold.

[2] The interpretation of figures like DeRose's implies a
theoretical problem of considerable importance. In this
connection we shall just point out that such figures always
presuppose a description; the percentages cited are thus no
raw facts about language.

[3] The term derivation does not, of course, carry any dia-
chronic overtones in the present connection.

[4] For an assortment of definitions and a good discussion of
traditional views, see Pennanen (1971:17-25).

[5] The description of (6)-(7) is particularly suspect, as it assumes lexical entries <INFINITIVE> and <IMPERATIVE> which are never realized phonologically and only appear with an overt element. Conversely, we also seem to be saying that <cut:V> can only be realized when accompanied with such a phantom element.

[6] Categorizing a verbalizing suffix as V etc. is adopted from Lieber (1981).

References

Bloch, Bernard 1947. English verb inflection. Lg. 23, 399-418. Also in Joos (1968, ed., 243-54).

Bloomfield, Leonard 1935. Language. Allen & Unwin: London.

Chomsky, Noam 1965. Aspects of the Theory of Syntax. The MIT Press: Cambridge, Mass.

------------- 1970. Remarks on Nominalization. In Roderick Jacobs and Peter Rosenbaum, eds., Readings in English Transformational Grammar. Ginn and Co: Waltham, Mass., 184-221.

Crystal, David 1967. English. Lingua (Word Classes) 17, 24-56.

DeRose, Steven 1988. Grammatical Category Disambiguation by Statistical Optimization. Computational Linguistics 14, 31-9.

Gleason, H.A. 1969. An Introduction to Descriptive Linguistics, Revised Edition. Holt, Rinehart and Winston: London.

Greenberg, Joseph H. 1963. Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In Greenberg, ed., Universals of Language. The MIT Press: Cambridge, Mass., 58-90.

Haas, W. 1957. Zero in Linguistic Description. In Studies in Linguistic Analysis. Special volume of the Philological Society. Blackwell: Oxford, 33-53.

Harris, Zellig S. 1946. From morpheme to utterance. _Lg._ 22, 161-83. Also in Joos (1968, ed., 142-53).

Horrocks, Geoffrey 1987. _Generative Grammar_. Longman: London.

Huddleston, Rodney 1984. _Introduction to the Grammar of English_. Cambridge University Press: Cambridge.

Hudson, Richard 1984. _Word Grammar_. Blackwell: Oxford.

Jackendoff, Ray 1975. Morphological and semantic regularities in the lexicon. _Lg._ 51, 639-71.

Jespersen, Otto 1942. _A Modern English Grammar. Part VI: Morphology_. Munksgaard: Copenhagen.

Joos, Martin 1968, ed. _Readings in Linguistics_, Vol. I. The University of Chicago Press: Chicago.

Koskenniemi, Kimmo 1983. _Two-Level Morphology: A General Computational Model for Word-Form Recognition and Production_. Publications of the Department of General Linguistics, 11. University of Helsinki.

Lieber, Rochelle 1981. Morphological conversion within a restrictive theory of the lexicon. In M. Moortgat, H. v. d. Hulst, T. Hoekstra, eds., _The Scope of Lexical Rules_. Foris: Dordrecht, 161-200.

Lyons, John 1977. _Semantics_, Vol 2. Cambridge University Press: Cambridge.

Marchand, Hans 1969. _The Categories and Types of Present-Day English Word-Formation_. Second edition. Verlag C.H.Beck: Munchen.

Nida, Eugen A. 1949. _Morphology. The Descriptive Analysis of Words_. University of Michigan Publications: Ann Arbor.

Nyman, Martti 1989. Combinability and situations: Is the Saussurean sign in need of augmentation? _The 1989 Yearbook of the Linguistic Association of Finland_. Helsinki, 31-52.

Pennanen, Esko V. 1971. Conversion and Zero-Derivation in English. _Acta Universitatis Tamperensis_, Ser. A, Vol. 40: Tampere.

Quirk, Randolph & Sidney Greenbaum & Geoffrey Leech & Jan Svartvik 1985. _A Comprehensive Grammar of the English Language_. Longman: London.