

[WLG]

WIENER LINGUISTISCHE GAZETTE

Contentives: a lexical supercategory above nouns and verbs

Hedde Zeijlstra

Sonderdruck aus: *Wiener Linguistische Gazette* (WLG) 82 (2017): 313–321

Themenheft 11-11-17. *Festschrift für Martin Prinzhorn*
Hg. v. Clemens Mayr und Edwin Williams

Universität Wien · Institut für Sprachwissenschaft · 2017

Eigentümer, Herausgeber und Verleger:

Universität Wien, Institut für Sprachwissenschaft
Sensengasse 3a
1090 Wien
Österreich

Redaktion: Christian Bendl, Mi-Cha Flubacher (Angewandte Sprachwissenschaft),
Stefan Schumacher (Allgemeine und Historische Sprachwissenschaft)

Kontakt: wlg@univie.ac.at

Homepage: <http://wlg.univie.ac.at>

ISSN: 2224-1876

NBN: [BL078,1063](https://nbn-resolving.org/urn:nbn:de:hbz:5:1-63864-p0078-1063)

Die *Wiener Linguistische Gazette* erscheint in loser Folge im Open-Access-Format.
Alle Ausgaben ab Nr. 72 (2005) sind online verfügbar.

Contentives: a lexical supercategory above nouns and verbs*

Hedde Zeijlstra

University of Göttingen

1. The question

It is a well-known fact that languages differ cross-linguistically with respect to what grammatical categories, or parts of speech, they exhibit. For instance, some languages exhibit (in)definite articles, whereas other languages lack them, and similarly, some languages lack adpositions, whereas other languages do not. This gives rise to the question as to what constrains the range of variation with respect to what grammatical categories languages may employ, and, more specifically, whether there are grammatical categories that can be attested in every language.

Traditionally it has been assumed that languages minimally distinguish nouns and verbs (cf. Baker 2003, 2008, Borer 2003, Croft 2003, 2005, 2009, Greenberg 1963, Halle & Marantz 1993, Pinker & Bloom 1990, Whaley 1997, a.o.). For some of these scholars, this universal noun-verb distinction is directly given by UG. However, a number of languages cast doubt on this assumption, as, at least overtly, such languages do not show any morpho-syntactic noun-verb distinction.

For instance, in a language like Samoan, all content words can systematically be used both verbally and nominally. For example, Samoan *alu* in (1) may either mean ‘to go’ or ‘(the) going’, depending on the grammatical context: combined with a tense marker it obtains a verbal reading ‘to go’; combined with an article, it yields a nominal reading ‘(the) going’ (cf. Mosel & Hovdhaugen 1992; Don & Van Lier 2013), as is illustrated below.

- (1) a. E alu le pasi I Apia. Samoan
 PRES go the bus to Apia (Don & Van Lier 2013)
 ‘The bus goes to Apia.’

*This paper has been presented at Olinco 2016. I thank the organizers and the audience for helpful feedback. Of course, many thanks, not so much for this paper, but for everything he did for the field, including training an impressive cohort of excellent linguists and making Vienna a wonderful place for linguistics, go to Martin Prinzhorn.

- b. Le alu o le pasi I Apia
 the go of the bus to Apia
 ‘the going of the bus to Apia’

Similar claims have been made for Mundari (Hengeveld and Rijkhoff 2005), Kharia (Peterson 2006) and Riau Indonesian (Gil 2013a,b). In Mundari and Kharia, just as in Samoan, content verbs may be used both nominally and verbally. For instance, in Mundari nominally used *buru* means ‘mountain’ and verbally used *buru* means ‘to heap up’ (2); in Kharia nominally used *lebu* means ‘man’ and verbally used *lebu* ‘to become a man’ (3).

- (2) a. Buru=ko bai-ke-d-a Mundari
 Mountain=3PL.S make-COMPL-TR-IND (Evans & Osada 2005)
 ‘They made the mountain.’
- b. Saan=ko buru-ke-d-a
 Firewood=3PL.S mountain-COMPL-TR-IND
 ‘The heaped up the firewood.’
- (3) a. Lebu del=ki Kharia
 Man came.MV.PST (Peterson 2006)
 ‘The man came.’
- b. Baghwan lebu=ki
 God man.MV.PST
 ‘God became a man.’

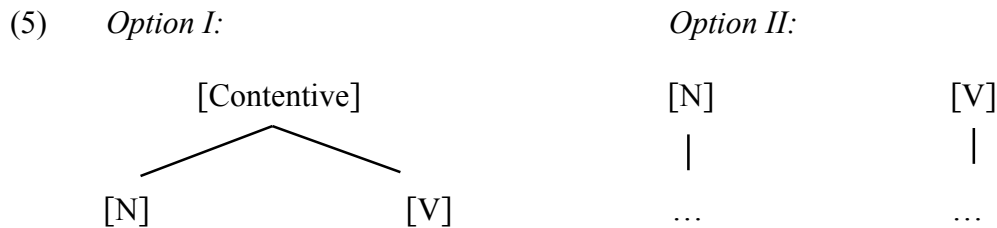
And for Riau Indonesian, Gil (2013a,b) has claimed that the syntactic distribution of any thing-denoting or action-denoting word is the same. E.g., both *abang* (‘(elder) brother’) and *kencing* (‘to pee’) can be modified by a demonstrative (as shown in (4) below), and may also appear in existential constructions, form the complements of adpositions, or combine with topic markers.

- (4) a. Abang in-i Riau Indonesian
 Elder.brother DEM-PROX (Gil 2013b)
 ‘that brother/man’
- b. Ter-kencing in-i
 Non_AG.pee DEM-PROX
 ‘to pee’

The question is thus whether the examples in Mundari, Kharia, Samoan and Riau Indonesian (and any other language that exhibits the same pattern) form counterexamples to the claimed universal noun-verb distinction, or whether these languages nevertheless underlyingly exhibit distinct nouns and verbs.

2. Contentives

As syntactic categories are reflections of categorial features, the question phrased above amounts to wondering whether nominal and verbal features ([N] and [V]) are sisters in a feature hierarchy. Is there a superfeature, that one can dub [Contentive], that immediately dominates [N] and [V], or are [N] and [V] the top nodes of a featural hierarchy of their own? The two options are illustrated in (5).



Hengeveld (1992, 2005) and Hengeveld & Rijkhof (2005) have argued that languages that lack a morpho-syntactic noun-verb distinction exhibit a supercategory dubbed *contentives*. In terms of featural hierarchies they (implicitly) hypothesize Option I in (5). The same conclusion has been reached by Mosel & Hovdhaugen (1992), who have also argued that languages such as Samoan lack distinct nouns and verbs and exhibit a single lexical supercategory instead.

By contrast, Croft (2005), among others, has argued that in this type of languages such nouns and verbs are actually homophonous: in (1) there are two instances of *alu*, a noun *alu* ‘(the) going’ and a verb *alu* ‘to go’. The central argument for postulating a noun-verb distinction in languages where there are no visible morpho-syntactic differences between nouns and verbs, and thus for denying the existence of contentives, is that the meanings of the verbal and nominal usages of such alleged contentives do not follow compositionally. For instance, Samoan *tusi* means ‘to write’, ‘letter’ and ‘book’. Similarly, Samoan *fana* means ‘to shoot’ and ‘gun’, *gaoi* ‘to steal’ and ‘thief’, and *eklaesia* ‘to go to church’ and ‘church member’. It would be very hard to come up with a fully compositional analysis purely on the basis of the linguistic environment that can derive these meanings (and these meanings only) from a single semantic core. Evans & Osafa (2005) for this and other reasons assume that languages that seem to exhibit contentives (Mundari is their example) actually involve zero-derivation and do not exhibit a lexical supercategory ‘contentives’.

In order to address these problems, Hengeveld & Rijkhoff (2005) have argued that in such flexible languages interpretation does not have to proceed in a strictly compositional fashion. For them, the underlying semantics is vague, so that both readings can come about, even though it is not predictable which meanings must be yielded. Contentives only have some flexible core, and the more specific meaning has to come about contextually.

The existence of contentives thus seems to rely on the assumption that languages may allow non-compositional interpretations. However, this restricted view on compositionality is at complete odds with its original motivation in Frege 1892 and Janssen 1986. Giving up compositionality immediately introduces the virtually unsolvable question as to why sentences in all other languages must be subject to

compositional interpretation. However, if indeed for contentives semantic compositionality is required, the question still arises as to whether the absence of fully predictable meanings is really an argument against the existence of contentives.

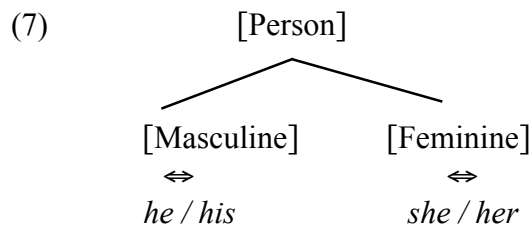
3. Proposal

In this paper, I pursue a novel, asymmetric approach. So far it has always been assumed that if a language exhibits two major lexical categories, they must always be nouns and verbs. Thus, if Samoan has two different lexical items *tusi* ('to write' and 'writing/letter/book'), one must be a noun *tusi* (meaning 'writing/letter/book') and the other a verb *tusi* (meaning 'to write'). No one, as of yet, has defended the view that some languages may have asymmetric categorical relations: languages having contentives and nouns, but lacking verbs, or languages having contentives and verbs, but lacking nouns. But nothing principled excludes such languages. If such contentives exist, they still must be taken to be some kind of supercategory above nouns and verbs as in Option I in (5).

As is known from the literature on morphological and semantic markedness, such asymmetries exist and can be diagnosed. Sauerland (2008), among others, has shown that semantic feature markedness is reflected by entailment relations (see also Heim 2008, Zeijlstra 2015). To see this, in a language like English, where feminine and masculine gender are not marked with respect to each other, both masculine and feminine pronouns have a gender-specific reference:

- (6) a. Everybody who lost his credit card must report it
→ masculine referential inference
- b. Everybody who lost her credit card must report it
→ feminine referential inference

In English, [masculine] and [feminine] must be daughters of some [person] feature; [masculine] and [feminine] do not entail each other:



In a language like Dutch things are, however, different.

- (8) a. Iedereen die zijn creditcard verloren is, moet dat melden
'Everybody who lost his credit card must report it.'
→ no gender-specific referential inference

- b. Iedereen die haar creditcard verloren is, moet dat melden
 ‘Everybody who lost her credit card must report it.’
 → feminine referential inference

Here the masculine and feminine sentences stand in an asymmetrical entailment relation. Following Zeijlstra (2015), Dutch lacks a feature [Masculine] and what looks like masculine morphology is actually the realization of a gender-unspecific [person] feature. This feature has only one daughter: [Feminine].

- (9) [Person] \Leftrightarrow *hij* (‘he’) / *zijn* (‘his’)
 |
 [Feminine] \Leftrightarrow *zij* (‘she’) / *haar* (‘her’)

Similar featural (a)symmetries have been proposed for 1st-2nd person distinctions, singular-plural distinctions, past-present tense distinctions, mass-count distinctions, comparative-superlative distinctions, etc. Languages appear to differ with respect to whether a particular categorial opposition underlies a featural sisterhood or mother-daughter relation (see Zeijlstra 2015 for more discussion and examples).

Feature hierarchies, like the ones above have not only been proposed for semantic features but also for morpho-syntactic features. For instance, person/number/gender features on verbs, which are semantically inactive, also stand in the same kinds of feature hierarchies as semantic features. Morpho-syntactic sub- and super-features are also hierarchically ordered. Consequently, in every language that distinguishes two opposite lexical categories, the question arises as to whether these categories are the realizations of the features of two sister nodes ([N] and [V]) or whether, they realize the mother node (the superfeature [Contentive]) and one daughter node [N]/[V], as in (10a) or (10b):

- (10) a. [Contentive] b. [Contentive]
 | |
 [N] [V]

In the first case, option I in (5), nouns and verbs are marked with respect to each other; in the second case, (10a) or (10b), the noun or verb would be marked with respect to the contentive that can be used both verbally and nominally.

4. Asymmetric meaning relations

Now, if a language does not exploit nouns and verbs, but, say, nouns and contentives, markedness effects as discussed above predict asymmetric meaning relations between the two. A contentive, being underspecified for being a noun or being a verb, should be able to appear both in nominal and in verbal morpho-syntactic contexts. By contrast, a noun,

by definition, may only appear in nominal morpho-syntactic contexts. Consequently, if the noun and the contentive have different meanings, we expect that in such a case, the meaning that is associated with the contentive should be available both in nominal and in verbal morpho-syntactic contexts (as the contentive can be used in both types of contexts), but that the meaning that is associated with the noun is available in nominal morpho-syntactic contexts only. *Mutatis mutandis*, the same applies if a language exploits only contentives and verbs.

This prediction is indeed borne out. Closer observation of the data presented in Mosel & Hovdhaugen 1992 and Don & Van Lier 2013 reveals that Samoan *tusi*, as well as a number of other examples, behave exactly like that. *Tusi* has a much richer nominal usage, varying from ‘(the) writing’ to ‘letter/book’, than a verbal usage: *tusi* in a verbal context can mean ‘to write’, but never ‘to be a letter/book’. This shows that there is a noun *tusi* meaning ‘letter/book’, and a contentive *tusi* that means ‘to write’ (in verbal contexts) and ‘the writing’ (in nominal contexts). The noun *tusi* (meaning ‘letter/book’) can only be used in morphologically nominal contexts; it can never appear in contexts that are morphologically verbal. By contrast, the contentive *tusi* (meaning ‘write/writing’) has no categorical restrictions and can therefore be used both in morphologically nominal and morphologically verbal contexts. The same pattern can be attested with, for instance, contentive *fana*, which means ‘to shoot’ and ‘(the) shooting’, and nominal *fana*, which means ‘the gun’. The reverse pattern has not been attested in Samoan: cases where the verbal usage would be semantically richer than the nominal usage.

These facts also extend to other languages. In Kharia, the word *bui* means ‘to keep’ or ‘(the) keeping’ (cf. Peterson 2006) and can be used both verbally and nominally with this meaning. But, overtly derived *bu-nu-i*, meaning ‘pig’, is only attested in nominal contexts. No examples of *bu-nu-i* meaning ‘to be a pig’ are attested (cf. Don & Van Lier 2013).

Hence, even though Samoan and Kharia seem to exhibit two different lexical categories (which solves the compositionality problem that would otherwise arise), it falsifies the claim that every language has nouns and verbs and confirms the claim that languages may exhibit contentives.

One caveat must be made, though. Under this perspective, categorial features like [N], [V] and [Cont] do not bring in any meaning contribution themselves. What features like [N] and [V] do is only restrict the grammatical distribution of their carriers (to nominal and verbal grammatical contexts respectively). The noun *tusi* in Samoan is a different word than the homophonous contentive *tusi*. If both *tusi*’s shared the same categorial feature [Cont], their grammatical distributions would be alike, and each word *tusi*, with its own meaning, could be used in each grammatical context. Then *tusi* would be predicted to be able to be used as ‘(the) letter’ and ‘(the) writing’ in nominal contexts, and ‘to be a letter’ and ‘to write’ in verbal contexts, contrary to fact.

In this sense, the meaning oppositions between the two *tusi*’s and, for instance, Dutch masculine and feminine pronouns, are different. In Dutch, a masculine pronoun that is unbound or does not appear in ignorance contexts, still receives a masculine interpretation, due to pragmatic blocking (arguably, Maximize Presupposition, cf. Heim 2008, Sauerland 2008 and references therein). But for examples like *tusi* this is not the case; the contentive *tusi* (meaning ‘write/writing’) does not further narrow down its meaning because of competition with the noun *tusi* (meaning ‘letter/book’). The reason is

simply that features like [V], [N] and [Cont] do not stand in an entailment or other semantic relation; these features only determine morphosyntactic distributions

5. Conclusions

The meaning observations for Samoan elements that may appear in both verbal and nominal morphosyntactic contexts show an asymmetry that is best explained by assuming homophony between a noun with one particular meaning and a contentive with another, crucially unpredictable, meaning. This, in turn, provides evidence for a lexical supercategory, contentives, above nouns and verbs, and thus for a superfeature [Contentive], that in a feature hierarchy immediately dominates [N] and [V].

At the same time, many questions emerge. Below I spell out some of those, but the list is of course far from exhaustive.

First, if [Contentive] is indeed a superfeature above [V] and [N], then this feature should also be part of the grammar of languages that clearly exhibit nouns and verbs. This would be evidenced by grammatical phenomena that are sensitive to the presence of a feature [Contentive]. One such phenomenon could be PP-modifiability. It is well known that PPs may modify nouns, verbs and predicatively used adjectives, but not DPs or attributively used adjectives:

- (11) a. wine from Austria
 b. living in Austria
 c. The world is afraid of Austria.
 d. *Martin from Austria (not intended as a single name)
 e. *the afraid of Austria country

If one were to define the types of elements that PPs may be the direct complement of modify (or that PPs may right adjoin to), one could argue that these are all elements that carry [Contentive], namely nouns, verbs and predicates (cf. Escribano 2004, Zeijlstra 2016).

A second, perhaps more pressing question, concerns the notion of morphological roots. Roots are generally assumed to be acategorial, as they are lexically underspecified for being nominal or verbal. But one could hypothesize that what acategorial roots are, is actually what contentives are: elements that belong to a lexical category without further specification. Hence, one may wonder whether roots are really category-less, or whether they are rather elements with a higher categorial feature [Contentive]. A root without any categorial feature, that becomes nominal or verbal due the morphology it attaches with, can be also thought of as a contentive that is further specified/valued by either verbal or nominal (sub)features.

Assuming that roots are actually contentives, would solve one problem for current morpho-syntactic theory, namely the fact that acategorial roots may merge with formal features (and therefore instantiate merger between elements that share no formal property). Allowing the syntactic operation Merge to apply to two elements that carry one or more formal features and to two elements of which only one carries a formal feature and the other does not (as would be the case when roots undergo merger), essentially

boils down to defining two different operations Merge: one for merger with roots, and one for merger with non-roots. Replacing roots by contentives would restrict syntactic Merge to elements that carry formal features, and thus only require one operation Merge.

In this small paper I do not pretend to have answers to such questions. However, if the reasoning behind the argument that there is indeed a supercategory contentive above nouns and verbs is correct, a novel opening towards addressing these questions can be pursued.

References

- Baker, Mark. 2003. *Lexical categories: Verbs, nouns and adjectives*. Cambridge: Cambridge University Press.
- Baker, Mark. 2008. *The Syntax of agreement and concord*. Cambridge: Cambridge University Press.
- Borer, Hagit. 2003. Exo-skeletal vs. endo-skeletal explanations: Syntactic projections in the lexicon. In *The nature of explanation in linguistic theory*, ed. John Moore & Maria Polinsky, 31-67. Stanford, CA: CSLI publications.
- Croft, William. 2003. *Typology and universals*. Cambridge: Cambridge University Press.
- Croft, William. 2005. Word classes, parts of speech, and syntactic argumentation. *Linguistic Typology* 9: 431-441.
- Croft, William. 2009. Syntax is more diverse, and evolutionary linguistics is already here. *Behavioral and Brain Sciences* 32: 453-454.
- Don Jan, & Eva van Lier. 2013. Categorization and derivation in flexible and differentiated languages. In *Flexible word classes: a typological study of underspecified parts-of-speech*, ed. Jan. Rijkhoff & Eva van Lier, 56-88. Oxford: Oxford University Press.
- Escribano, José. 2004. Head-final effects and the nature of modification. *Journal of Linguistics* 40: 1-43.
- Evans, Nicholas, & Toshiki Osada 2005. Mundari: the myth of a language without word classes. *Linguistic typology* 9: 351-390.
- Frege, Gottlob. 1892. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100: 25-50.
- Gil, David. 2013a. Word order without syntactic categories: how Riau Indonesian does it. In *Verb first: on the syntax of verb-initial languages*, ed. Andrew Carnie, Sheila Ann Dooley, & Heidi Harley, 243-263. Amsterdam: Benjamins.
- Gil, David. 2013b. Riau Indonesian: A language without nouns and verbs. In *Flexible word classes: a typological study of underspecified parts-of-speech*, ed. Jan Rijkhoff & Eva van Lier, 89-113. Oxford: Oxford University Press.
- Greenberg, Joseph. 1963. *The Languages of Africa*. Bloomington, IN: Indiana University Press.
- Halle, Morris, & Alec Marantz 1993. Distributed Morphology and the pieces of mnflection. In *The View from Building 20*, ed. Ken Hale & Samuel J. Keyser, 111-176. Cambridge, MA: The MIT Press.

- Heim, Irene. 2008. Features on bound pronouns. In *Phi theory. Phi-features across modules and interfaces*, ed. Daniel Harbour, David Adger, & Susana Béjar, 35–56. Oxford: Oxford University Press.
- Hengeveld, Kees. 1992. Parts of speech. In *Layered structure and reference in a functional perspective*, ed. Michael Fortescue, Peter Harder, & Lars Kristoffersen, 29–55. Amsterdam: Benjamins.
- Hengeveld, Kees. 2005. Parts of speech. In *Crucial readings in functional grammar*, ed. Matthew Anstey & J. Lachlan Mackenzie, 79–106. Berlin, New York, NY: Mouton de Gruyter. [Reprint of Hengeveld (1992).]
- Hengeveld, Kees, & Jan Rijkhoff 2005. Mundari as a flexible language. *Linguistic Typology* 9: 406–431.
- Janssen, Theo. 1986. *Foundations and applications of Montague Grammar*. Doctoral dissertation, University of Amsterdam, Amsterdam.
- Mosel, Ulrike, & Even Hovdhaugen 1992. *Samoan reference grammar*. Oslo: Scandinavian University Press.
- Peterson, John. 2006. *Kharia. A South Munda language. Volume I: Grammatical analysis*. Habilitation Thesis, Osnabrück University, Osnabrück.
- Pinker, Steven, & Paul Bloom 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13: 707–784.
- Sauerland, Uli. 2008. On the semantic markedness of phi-features. In *Phi theory. Phi-features across modules and interfaces*, ed. Daniel Harbour, David Adger, & Susana Béjar, 57–82. Oxford: Oxford University Press.
- Whaley, Lindsay. 1997. *Introduction to typology: The unity and diversity of language*. Thousand Oaks, CA: Sage.
- Zeijlstra, Hedde. 2015. Let's talk about you and me. *Journal of linguistics* 51: 465–500.
- Zeijlstra, Hedde. 2016. *Projection and feature percolation*. Ms., University of Göttingen.

Hedde Zeijlstra
hzeijls@uni-goettingen.de