# From the editors

**Miriam van Staden**
**Hedde Zeijlstra**

We are happy to introduce the second issue of our second volume of the ACLC Working Papers. After two years, the Working Papers are taking root in the ACLC, evidenced by the diversity of papers that have been submitted. Experimental phonetics, second language acquisition, functional and formal theories all find a place in the ACLC and in our Working Papers.

This year, Enoch Aboh resigned as one of the two editors-in-chief. On behalf of the entire editorial board, we would like to thank him for his input in initiating the Working Papers. Enoch has been replaced by Hedde Zeijlstra. Next year, there will be a further change in the editorial board as Miriam van Staden will be replaced by Judith Rispens.

What will not change is our endeavor to publish up-to-date and state-of-the-art research carried out at the Amsterdam Center for Language and Communication at the University of Amsterdam.

# ACLC Working Papers
# M. van Staden & H. Zeijlstra (eds)
# Volume 2, issue 2

# Integrated participle clauses From adverbial to complement*

**Hendrik De Smet**
*Fund for Scientific Research (FWO) – Flanders*

*In the course of the Modern and Present-Day period, participle clauses have in a variety of constructions come to be more closely integrated in the syntax of the matrix clauses they combine with, thus developing into a special type of complement-like slot fillers. Such participial complement clauses will be referred to here as Integrated Participle Clauses or IPCs. The following discussion first provides a synchronic characterisation of the constructions in question, highlighting the syntactic and semantic differences between IPCs and other clause types (section 1). The synchronic discussion is followed by a diachronic analysis that looks into the possible sources of IPCs (section 2). To complete the picture, the ontological question is addressed concerning the precise status of IPCs. Doing so, the findings of the two previous sections are situated against a more general background: it is proposed that, diachronically, IPCs can be seen as a natural diachronic development from adverbial to complement, while synchronically, they can alternatively be seen as an independent category, a loose set of unrelated constructions, or marginal members of the category of -ing-complements in English (section 3). The discussion ends with a number of concluding remarks concerning the various ways in which IPCs exemplify gradience in language (section 4).*

## 1 Synchronic characterisation

The italicised -*ing*-forms in (1) and the clauses of which they are the head illustrate the kind of constructions that will be considered here as IPCs:[1]

* The research reported in this paper has been made possible by a grant from the Fund for Scientific Research (FWO) – Flanders. The research was carried out at the Amsterdam Center for Language and Communication. I would especially like to thank Olga Fischer for reading and commenting on two earlier versions of this paper. Further, I also benefited from some of the comments of participants at the fourteenth ICEHL conference in Bergamo. I also gratefully acknowledge the comments by Harry Perridon, who reviewed my paper for the *ACLC Working Papers*.

[1] Data used in this paper have been collected from various corpora. Present-Day English material has been gathered primarily from the *Collins Cobuild Corpus* (CB) (containing about 56 million words), and occasionally from the *British National Corpus* (BNC), the *International*

(1) a. The receptionist is busy *filling* a fifth box. (CB)
b. I am tired *hearing* of the Duchess of Chiselhurst's ball (1899, CEN)
c. Mr Jones said because he was not being properly paid he had trouble *getting* a housing loan and feared he might lose his new home. (CB)

IPCs have received hardly any attention in the synchronic or diachronic literature. The type in (1a), headed by the adjective *busy*, has been recognised as an adjective complement construction by Quirk *et al.* (1985: 1230) and Huddleston & Pullum (2002: 1259). The fact that Quirk *et al.* give an additional example with the adjective *fortunate* ("we're *fortunate* having Aunt Mary as a babysitter", 1985: 1231) indicates that they consider the construction productive. However, no further description is given of the kinds of predicates that occur in this construction, nor is there any discussion of what distinguishes the construction from other construction types such as gerund clauses and adverbial participle clauses.

The non-nominal nature of the *-ing*-clauses in (1) is obvious from their inability to alternate with an NP (consider, for instance, **I am tired all this talk of the Duchess of Chiselhurst's ball*). From this it may be concluded that they are in all likelihood not gerunds in a strict sense and by that token differ from the clauses illustrated in (2).

(2) a. I think people are tired of *hearing* about it. (CB)
b. Along with the rest of his partners, he will have to weigh up whether to go public now, or risk *remaining* private. (CB)

Not being nominalisations, the IPCs in (1) (as their name implies) more closely resemble participle clauses, or PCs. At least, they share their non-nominal character with the adverbially used PCs in (3).

---

*Corpus of English – Great Britain* (ICE-GB), *Google*, and occasional printed sources. Historical data have been gathered primarily from the extended version of the *Corpus of Late Modern English texts* (CLMETEV) (the last section of which covers the period 1850-1920 and contains about 6.3 million words) and the *Corpus of English Novels* (CEN) (which covers the period 1881-1922 and contains about 25 million words) (for more information on these, see the website http://perswww.kuleuven.be/hendrik_desmet). Where necessary the historical data have been supplemented with data from the *Penn Parsed Corpus of Early Modern English* (PPCEME), the *Corpus of Early Modern English texts* (CEMET), the *Lampeter Corpus* (LC), the *Proceedings from the Old Bailey*, and the *Innsbruck Middle English Prose Corpus* (ICAMET).

(3) a. See how many words of four or more letters you can find *using* the letters above (CB)
b. Fishermen in Scotland have taken a tennis club to court, *claiming* that its floodlights are driving away the fish in an angling river. (CB)

Still, it can be argued that IPCs also differ from adverbially used PCs, because their relation to the matrix predicate is not one of adverbial modification but of complementation – in this respect IPCs are in fact more closely related to the gerund clauses in (2) than to the PCs in (3). The purpose of this section is to address the differences between gerund clauses, adverbial PCs and IPCs. Because the relation between IPCs and adverbial PCs is more intricate the primary focus is on mapping out in what ways IPCs and adverbial PCs diverge and in what respects they are similar – though later on gerundial and other complements are drawn into the discussion as well.

### *1.1 IPCs, adjuncts and disjuncts*

Looking at the contrast between IPCs and adverbial PCs, it should be pointed out from the outset that IPCs do not differ from all adverbial PCs to the same extent. For one thing, adverbial PCs can themselves be divided into adjuncts and disjuncts. This distinction, based originally on Greenbaum (1969: 15-25) and Quirk *et al*. (1985: 1070-3), is illustrated in (3) above, with (3a) exemplifying an adjunct, and (3b) presenting a disjunct. Adjuncts in particular share certain features with IPCs, as will become clear in the following discussion. Let us start, therefore, by contrasting IPCs, adjuncts, and disjuncts as three broad and somewhat idealised, maximally differentiated categories, ignoring for now the more subtle differences between the IPCs in more specific environments, as well as the more problematic cases where the distinction between IPC and adverbial participle are obscured (see section 1.2 below).

One set of characteristics distinguishes IPCs from disjuncts but not from adjuncts. Firstly, IPCs are not separated from the matrix clause by an intonational boundary or, in writing, a comma. In this they differ from disjunctively used PCs (cf. Kortmann 1991). Thus it is (intuitively) evident that adding an intonational break changes the semantic relation between the PC and its matrix clause. Compare example (4) below to (1b) above: in (4) the PC gives the reason why (the speaker assumes) the subject of the matrix clause must be tired, while in (1b) the PC specifies what the matrix subject is tired of.

(4) You must be tired, *wandering* about on the hills as you do! (1887, CEN)

While the difference between IPCs and disjuncts is clear, the absence of an intonational boundary does not distinguish IPCs from adjunctively used PCs, which can similarly form a single intonation unit with the matrix clause, as in (5).

(5) Bold, 31, now in Glenochil jail, claimed he hurt his knee *playing* at Perth prison. (CB)

Secondly, IPCs, like adjuncts but unlike disjuncts, can be the focus of negation and polarity questions in the matrix clause. Compare the possible interpretations of the IPC in (6), the adjunct in (7) and the disjunct in (8):

(6) a. Yet, for most people, the impression gained of Mr Hanley is that he has trouble *deciding* which shoe to put on first in the morning. (CB)
b. Mr Hanley doesn't have trouble deciding… ('Mr Hanley has trouble doing some things, but he does not have trouble deciding…')
c. Does Mr Hanley have trouble deciding… ? ('Mr Hanley has trouble doing some things, but does he have trouble deciding…?')

(7) a. Ian went to check the cars and found that they were missing their wipers and he went and told the hotel manager, who came out *looking* very worried. (CB)
b. The manager did not come out looking very worried. ('the manager came out but he didn't look very worried')
c. Did the manager come out looking very worried? ('the manager came out but did he look very worried?')

(8) a. Yesterday the group issued its strongest warning yet, *telling* foreigners to leave the country. (CB)
b. The group didn't issue its strongest warning yet, telling foreigners to leave the country (*'The group issued its strongest warning yet but did not tell foreigners to leave the country')
c. Did the group issue its strongest warning yet, telling foreigners to leave the country? (*'The group issued its strongest warning yet, but did it tell foreigners to leave the country?')

Thirdly, neither IPCs nor adjuncts can have explicit subjects – again in contrast to disjuncts, which do readily take their own subjects, as in (9a). Examples (9b-c) illustrate the effect of adding a subject to an IPC and adjunct; with the former the result is clearly ungrammatical, with the latter the result is acceptable only under a disjunct reading.

(9) a. It was only to be expected, he *being* thirty-five years older than me, but I can tell you expecting makes no difference. It's still an outrage to be left without a husband. (CB)
b. *Mr Hanley has trouble his wife deciding which shoe to put on first in the morning.
c. *The manager came out his assistant looking very worried.

Moreover, when they do not express their own subjects disjuncts can be controlled by the speaker/interlocutor, instead of a participant in the matrix clause, as in (10a), or the controller can be the matrix clause as a whole, as in (10b) (Kortmann 1991). Again, neither adjuncts nor IPCs have this possibility.[2]

(10) a. *Speaking* of which, why can't veterans just forgive and forget Japan over its treatment of allied prisoners of war so long ago? (CB)
b. In North Antrim at the last election, about 33,600 people voted for Unionist parties and 8,400 for nationalists, *indicating* that Catholics form about 20 per cent of its population. (CB)

A second, more important set of characteristics distinguishes IPCs from adverbial PCs in general – i.e. disjuncts as well as adjuncts – establishing them as a separate phenomenon. Firstly, omitting the IPC from the main clause will affect the meaning of the main clause predicate. On the one hand, the IPC restricts the meaning of the main clause predicate by specifying that the situation predicated in the main clause holds strictly in relation to the situation denoted by the PC, as in (11). As a result, omission of the PC broadens the semantic scope of the predicate. On the other, the IPC maps onto a semantic role evoked by the predicate and its omission brings about a shift from a transitive to an intransitive reading that may affect the lexical semantics of the predicate as such. This is particularly clear in (12): Without IPC, *tired* in (12a) would by default designate a state of physical (and possibly mental) exhaustion, rather than a feeling of weariness with relation to one specific activity. *Happy* in (12b) without IPC would designate a state of positive psychological excitement, but with IPC designates an attitudinal relation between a sensing subject and an object of emotional judgement. In contrast to all this, omission of a PC has no effect on the semantics of the matrix clause when the PC functions as an adjunct or disjunct, as can be tested by leaving out the PCs in (13) and (14) respectively.

[2] Incidentally, adjuncts can occasionally be controlled by the subject of a previous clause, as in the following example: *Third party developers were reluctant to release applications for Windows, complaining that it was slow. This was certainly true* running *on the 6MHz ATs of the day* (ICE-GB).

(11) a. I was fed up *sitting* at the station doing nothing. (CB)
b. Thousands like us need help *finding* someone special. (CB)
(12) a. The day I say I'm tired *playing* for my country is the day I hang up my boots. (CB)
b. I wasn't happy *being* described as cute, but seeing as I got the part I didn't care. (CB)
(13) a. As ever he stormed away *refusing* to speak, along with his sulking team. (CB)
b. I shall make so much money *exploring* Africa I shan't know what to do with it. (1902, CLMETEV3)
(14) a. In a classroom in Farmington […] about a dozen farmers are lined up in desks, *looking* at charts of farm prices projected on a screen. (CB)
b. Cyril glanced at Amy, who averted her head, *putting* spoons into three saucers. (CLMETEV3)

Judging by the examples in (11) and especially (12), the semantics of IPC-constructions resemble those of complement constructions. The semantic relation between the complement and the matrix clause is typically determined by the semantics of the matrix clause predicate. More precisely, a relation is predicated between the matrix clause subject and the situation described by the complement. This relation – abstractly defined – consists in (real or potential) psychological or physical energy either exerted by the matrix clause subject on the situation of the complement or triggered by the situation of the complement in the subject.

Secondly, another indication of the higher degree of integration of IPCs in the matrix clause is the fact that IPCs allow wh-extraction (cf. Los 2005: 48-9, who, following Chomsky 1980, uses this argument to assert the argument status of *to*-infinitives in Old English). An example of wh-extraction is given in (15). As to adverbial PCs, both adjuncts and disjuncts resist this operation (albeit to different degrees; see below). This is illustrated in (16) and (17).

(15) It's difficult to get the actual legislatures to act, and therefore one has to activate the legislators to do something. And this is what we are busy *trying* to do, and we have been preparing for this for the last 18 months. (CB)
(16) a. Police late yesterday were still looking for the youth, who escaped on foot *wearing* a baseball cap and a false beard. (CB)
b. *The false beard the young man escaped on foot wearing.
(17) a. The operation was successful for the commandos who managed to push the enemy (infantry and armoured units) onto the back foot and

keep them at bay for three days, *taking* over the town of Wyndham, WA, in process. (CB)

b. *The town that the commandos kept the enemy at bay for three days, taking over in process.

Thirdly, IPCs can never be introduced by subordinating conjunctions such as *when*, *while*, *if* or *though*. Adverbial PCs allow this option, provided that the semantics of the inserted conjunction do not disagree with the meaning of the sentence. Compare the effect of inserting a conjunction before the IPC in (18) and the adverbial PC in (19).

(18) a. The old man, 65-year-old Anatole Pierre, is busy *digging* up the roots of a felled mahogany. (CB)

b. *The old man is busy while/when digging up the roots of a felled mahogany.

(19) a. Stan and his remaining survivor a Chinese man, had radio contact with the Allied forces who tried to send submarines in to rescue them but lost men *doing* that. (CB)

b. The Allied forces lost men while/when doing that.

Fourthly, adjuncts – though not disjuncts – can often be questioned by a wh-question using adverbial interrogative pronouns such as *how*, *when* or *why*, as in (20). IPCs, by contrast, generally resist questioning – compare (21a-b). If IPCs are questioned, the most natural interrogative pronoun to use is *what* – as is illustrated in (21c) – but notice that questioning then typically requires the addition of a preposition to the matrix clause predicate, which suggests that what is questioned in such examples is not the IPC but a semantically similar gerundial construction with preposition (e.g. *Jeff has had a lot of success in breeding Doves*).[3]

(20) a. We started the season *aiming* for the top four and a place in Europe. (CB)

b. How did we start the season?

(21) a. Jeff has had a lot of success *breeding* Doves, but for six years kept a par [sic] of Tambourines, without them making any attempt to nest. (CB)

[3] An alternative questioning strategy avoids the addition of a preposition and instead makes use of extraposition, keeping the IPC-slot filled with semantically 'empty' *doing* and using the question word *what* as its extraposed object, e.g. *what did Jeff have a lot of success doing?*

b. *How/when has Jeff had a lot of success?
c. What has Jeff had a lot of success in?

In order to summarise the discussion so far, the various characteristics of IPCs, adjuncts and disjuncts are repeated here in Table 1.

**Table 1**: Characteristics of IPCs, adjuncts and disjuncts.

| IPC | Adjunct | Disjunct | |
|---|---|---|---|
| + | + | - | i. PC is not separated from the matrix clause by an intonational boundary |
| + | + | - | ii. PC falls under the scope of negation, polarity questions and epistemic modals in the matrix clause |
| + | + | - | iii. PC has no explicit subject |
| + | - | - | iv. Omission of PC changes the semantics of the matrix clause predicate, broadening the scope of the predication |
| + | - | - | v. PC allows wh-extraction |
| + | - | - | vi. PC resists insertion of subordinating conjunctions |
| + | - | n/a | vii. PC cannot be questioned by *how* or other adverbial interrogative pronouns |

At this point, it is clear that IPCs behave rather more like complements than adverbial PCs. That is, IPCs and more common types of complement clauses show roughly the same behaviour under the various tests listed in Table 1. This is shown in the following examples where the tests distinguishing IPCs from adjuncts (iv-vii in Table 1) are applied to constructions with a gerundial and infinitival complement ((22) and (23) respectively). As is clear from these examples, there is little difference between IPCs and gerundial or infinitival complements in terms of the relation that obtains between the matrix clause and the dependent clause.

(22) a. She'll throw her toys around and will enjoy *making* a mess with her dinner. (CB)
b. *She'll throw her toys around and will enjoy.
c. The mess she will enjoy making.
d. *She'll throw her toys around and will enjoy while making a mess with her dinner.
e. *How does she enjoy?

(23) a. I guess I wanted *to be* a hero. (CB)
b. *I guess I wanted.

c. The hero that I wanted to be.
d. *I guess I wanted while/in order to be a hero.
e. *How did I want?

At the same time, it must be recognised that in at least some respects IPCs differ from the more typical gerundial and infinitival complement constructions. One difference is that other complement constructions often more strongly resist omission of the complement clause than IPC constructions, as the main clause predicate not just changes meaning, but in fact becomes ungrammatical (as in (22b) and (23b) above). Further, IPCs seem less inhospitable to the insertion of adverbial material between the complement-taking predicate and the complement clause – an operation which gerundial complements at least tend to resist – as is illustrated in (24).[4] Regular complement clauses standardly allow questioning by *what* unlike IPCs, which sometimes simply cannot be questioned by a wh-pronoun or at best require the addition of a preposition – compare (21) above with (25). Other complement types can have their own subject, as shown in (26). And, finally, because IPCs are not nominals, they resist some additional operations that are allowed by the more regular complement types, such as (pseudo-)clefting or pronominalisation, as in (27) and (28).

(24) a. Police departments don't change easily, and Williams had trouble early in his tenure *identifying* commanders who wanted to follow the community policing model. (CB)
b. In Bao Loc – a highlands hole-in-the-wall four hours north of Ho Chi Minh City, dozens of families were busy last month *attempting* to turn homes into mini-hotels. (CB)
c. *She enjoys every day making a mess with her dinner.
d. I wanted all my life to be a hero.

(25) a. What will she enjoy?
b. What did I want?

(26) a. Suzie's parents find it hard to enjoy her making a mess with her dinner.
b. I guess everybody wanted him to be a hero.

(27) a. What she'll enjoy is making a mess with her dinner.
b. What I wanted was to be a hero.

[4] The fact that *to*-infinitival complements less strongly resist the insertion of adverbial material between main verb and complement may reflect the lower degree of 'nouniness' of *to*-infinitives as compared to gerund clauses (Ross 1973), or may reflect the fact that, like IPCs, *to*-infinitival complements historically derive from adverbial constructions (see section 2.2).

c. *What Williams had trouble was identifying commanders who wanted to follow the community policing model.
d. *What dozens of families were busy was attempting to turn homes into mini-hotels.

(28) a. She'll enjoy it.
b. (?) I guess I wanted it.
c. *Williams had trouble it.
d. *Dozens of families were busy it.

Very much the same thing can be said about constructions where a gerund is introduced by a preposition, as in (29a). Arguably, these resemble IPCs somewhat more closely than gerunds, because it is often the same predicates that take IPCs that also occur in the prepositional pattern, without much appreciable difference in meaning. At the same time, the characteristics relating to the nominal nature of the gerund also apply to the patterns with preposition, and serve to distinguish IPCs from their prepositional variants. Tests demonstrating the similarities between IPCs and prepositional variants are applied in (29b-e); tests highlighting the differences in (29f-i)

(29) a. I am tired of *torturing* my family. (CB)
b. (?) I am tired.
c. The family that I'm tired of torturing.
d. *I am tired of while/when torturing my family.
e. *Why/how/when am I tired?
f. What am I tired of?
g. I am tired of Miss Whoopsy torturing my family.
h. What I am tired of is torturing my family.
i. I am tired of it.

It may be concluded that the differences between IPCs and adverbials are sufficiently convincing to treat IPCs as a distinct phenomenon, closely resembling other means of clausal complementation. Differences that exist between IPCs and other complement clauses mostly relate to the non-nominal status of IPCs.

### *1.2 Between adverbial and complement clause*

While the characteristics listed in Table 1 can be used to set IPCs apart from both adjuncts and disjuncts, they do not exhaust the differences that exist between various constructions and say as yet nothing about the possibility of intermediate category membership. With a view to refining syntactic description, it is therefore crucial also to examine the various clusters of related construc-

tions that may qualify – to various degrees – as IPCs. It will become clear from the discussion that not all of the tests used to distinguish IPCs from adverbial PCs yield clear-cut results, and that not all constructions are equally clear examples of IPCs. In light of this, it is tempting to view IPCs as a gradient category, with some prototypical IPC constructions possessing all defining features of the class, and other constructions somewhat more closely resembling constructions with an adverbial PC (see e.g. Taylor 1998 on prototype effects in syntax) (but see section 3 on the categorial status of IPCs).

### *1.2.1 Adjective + IPC*

At least superficially, IPC-constructions can be classified by the syntactic structure of their main clause predicate. Thus, as a starting point, two major groups of (sometimes seemingly) similar constructions can be distinguished. The first group consists of predicatively used adjectives combining with an IPC that functions as postmodifier or complement to the adjective. Amply illustrated above, the construction type under consideration is exemplified once more in (30).

(30) New Man, that sociological phenomenon said to treat women as his equal and who is happy *sharing* domestic chores, was pronounced dead yesterday. (CB)

Semantically, the adjectives used in constructions of this kind fall into a number of sub-categories. The first set of adjectives express an emotive relation between the subject of the matrix clause and the situation designated by the PC, specifying how the former is emotionally affected by the latter. These adjectives include *bored*, *comfortable*, *fed up*, *happy* (as in (30)), *tired*, *uncomfortable* and *unhappy*. A second group of adjectives does not denote a psychological state, but an external judgement attached to the subject in respect of the situation denoted by the PC. Such adjectives include *right*, as in (31a), *better off*, *brave*, and *lucky*. A third group of adjectives express a relation of active occupation of the main clause subject in the situation denoted by the PC. These include most notably the adjective *busy*, as in (18a) above, but also, more marginally, *employed*, *engaged* and *occupied*. Potential members of this group are also the expressions *be gone*, *be off* and *be out*, as illustrated in (31b). Semantically related to the adjectives of active occupation is a fourth group denoting the manner or degree to which the matrix clause subject is advancing or has advanced in realising the situation denoted by the PC. Adjectives of this kind are *late*, as illustrated in (31c), *quick* and *slow*; and the group might be further expanded with the expressions *be done* and *be finished*, as illustrated in (31d-e).

(31) a. The Supreme Court in a five to four decision declared that the officer was right in arresting her; he was right in putting her in handcuffs; he was right *taking* her into custody, taking her to jail; and it was right to force her to post a bail of more than $300. (Google)
b. [G]enerally I was out *shoveling* long before my 'young lady' had her nightcap off. (1869, CLMETEV3)
c. What happens if I'm late *paying* my VAT? (CB)
d. Karen came through the door, lugging the bulky file. "Schultz is done *burning* copies," she said as she strode to Winters's desk and plopped the bundle down in front of him. (CB)
e. They must be finished *painting* by now. (CB)

Some of the adjectival constructions under discussion respect all the defining characteristics of IPCs, but as is to be expected, this is by no means always the case; indeed, the superficially similar structures exemplified in (30-31) reveal various differences when considered more closely. The most straightforward members of the category of IPCs are the constructions with emotive adjectives: omission of the PC has a clear semantic effect on the matrix clause predicate, wh-extraction is invariably allowed, and PCs can be questioned by *what* with relative ease (provided a preposition is added to the matrix clause predicate). Moreover, unlike most IPCs, these emotive adjectives resist the insertion of adverbial material between the adjective and the IPCs – when such material is inserted, as in (32), the PC automatically receives an adverbial interpretation (compare this to the examples in (24) above). In light of this, it is fair to treat the IPCs with emotive adjectives as genuine adjectival complements.

(32) a. They were never tired *telling* me. (1913, CEN) ('they were never tired of telling me')
b. They were tired yesterday telling me. ('they were tired yesterday when / as a result of telling me')

A number of the non-emotive adjectives, viz. *busy*, *late*, *done* and *finished*, also take IPCs approaching genuine complements, even if the semantic effect of omitting the PC is sometimes less dramatic (e.g. if a person is busy making their bed, it is fair to say they are busy, while it is not the case that a person who is happy sleeping on the coach is simply happy). Note further that the expression *be done* shows a remarkable alternation between the copula *be* and the perfect auxiliary *have* – a pattern obviously not shared by any of the other adjectives.[5]

[5] Note that if *finished* shows the same alternation, this remains invisible since the pattern with *have* is indistinguishable from the verb complement construction with gerund clause.

More problematic are the adjectives of the *right*-type. With these adjectives the semantic effect of omitting the PC is again less outspoken. Further, the adjectives *right, lucky* and *brave* differ from the above adjectives in that they occur in an alternative though semantically reasonably similar structural configuration in which the situation otherwise expressed by the PC is now expressed in a *that*-clause or *to*-infinitive and takes the syntactic function of extraposed subject – compare the examples in (33) (see also (31a) above, where the two uses occur together in a single sentence).

(33) a. Listen, we were really lucky *getting* home at all. (BNC)
b. Listen, it was really lucky we got home at all.

As to the adjectival expression *be better off*, this pattern resembles *right, brave* and *lucky* in being commentative in meaning, but *better off* does not occur in the construction illustrated in (33b). What is more, it is ill-behaved in allowing *how*-questions instead of *what*-questions and licensing the insertion of the subordinating conjunctions *when* or *if*, suggesting that although they license wh-extraction, the PCs following *better off* are really adverbial, as is further supported by the possibility of replacing the PC by a PP or adverb without a change in meaning (cp. *I'd be better off living under the bridges of Paris* and *I'd be better off there*).[6]

Finally, the same possibility of replacing the PC by a functionally similar adverb seems to be supported by the expressions *be gone*, *be out*, and *be off* (cp. *He is out working* and *he is out in the garden*), but unlike *better off* they resist questioning by an adverbial wh-pronoun. In this light, these expressions might be more suitably analysed as being akin to catenative constructions with *lie, sit* or *stand* and a PC, as in (34). As the latter are characterised by the reduced semantic prominence of the main verb (cf. Quirk *et al.* 1985: 506), a shared analysis for the catenatives and *be gone*, *be out* and *be off* ties in with the intuition that with all of these constructions omission of the PC is not impossible but strongly increases the prominence of the matrix clause predicate.[7]

(34) He stood *breathing* gusts of vapor into the snowflakes that flitted about his face and clogged his eyelids. (CB)

[6] The same possibility of inserting *when* or *if* applies to the adjectives *occupied*, *engaged* and *employed* when preceded by the adverb *better*.
[7] For *be gone* this statement may have to be qualified: *John is gone fishing* implies that John went out purposely in order to go fishing – an implication of intentionality that is missing with *John is gone*.

### *1.2.2 Light verb + IPC*

The second major group of interrelated constructions involve an IPC attached to a light verb (typically *have* but sometimes also *find* or *experience*) and its semantically 'heavy' NP. The IPC can be interpreted as a complement to the whole light verb idiom or as a postmodifier or complement to only the NP. This construction type is illustrated in (35):

(35) a. I give advice to people who are having difficulties *getting* a job, […]. (CB)
b. […] to believe in the possibility of events that I have a hard time *believing* will come to pass. (*My name is red*, p.233)

Again, different semantic sub-classes can be recognised. The most prominent group is constituted by light verb constructions with the NPs (*no*) *difficulty*(/*ies*), *problem*(*s*) and *trouble*. Some semantically related but less frequent expressions also take IPCs, such as *have* (*no*) *success* / *hesitation*, *have a tough job* or *have a hard time*. Of these, especially the constructions with *difficulty*(/*ies*), *problem*(*s*), *trouble* and *success* conform to all characteristics of IPCs and can be fairly naturally interpreted as clausal complement constructions, as they establish a relation between the matrix clause subject and a PC, specifying to what degree the subject is successful in realising the situation denoted by the PC. It is to be noted, however, that the possibility of using the NP along with the IPC without a light verb, as in (36), supports an alternative analysis of the IPC as a postmodifier or complement to the NP.

(36) a. I know the point you made about the difficulties *finding* evidence but you know the trail doesn't go cold just because forty fifty years have passed. (CB)
b. Bailey White is on summer vacation from her job *teaching* first grade in south Georgia. (CB)

Other light verb constructions are built on the NPs *business* and *right*, as exemplified in (37). Characteristic about these highly idiomatic constructions is that they always contain a negative element, typically *no*.

(37) a. The state has no right *telling* the people what they can and can't do with their own body. (CB)
b. But they can only search the parts of the house that a person could be hiding in. They have no business *looking* in a one foot square box for a 6ft. 20stone man. (CB)

Besides the obligatory negative element, these constructions differ somewhat from those of the *difficulty*-type in that they strongly resist questioning by any question word. Also, it seems impossible to use the sequence of NP and IPC without the light verb, so the IPC is unambiguously a complement to the light verb idiom rather than a complement or postmodifier to the NP. The meaning of the construction certainly fits in comfortably with a complement reading, as a modal relation of (external) permission is predicated of a subject in relation to the activity denoted by the IPC. A final characteristic is that the idiomatic meaning of the construction is lost without the IPC, and that under omission of the IPC the expressions feel incomplete (unless its content is contextually given).

### *1.2.3 Other IPC-constructions*

These two major sets of constructions – IPCs with predicative adjectives and IPCs with light verb idioms – do not cover all the constructions showing IPCs. Thus, (38) lists and illustrates a number of less clearly related constructions that appear nonetheless also to involve a complement-like PC, judging by the tests defining IPCs. These include mostly IPCs with other light verbs or verbal idioms – *take one's time* (38a), *take turns* (38b), *have one's hands full* (38c), *need help* (38d) – but also with some verbs – *hesitate* (38e), *assist* (38f) and *succeed* (38g).

(38) a. The conman turns up with a woman said to be his mother, drops the names of fancy hotels, flashes his mobile phone and asks to test drive expensive motor cars. The problem is, according to the local newsletter Motor Industry News, he takes his time *bringing* them back. (CB)
b. We took turns *opening* the refrigerator door and *hoping*, but no matter how many times we looked inside, the contents never changed. (CB)
c. Ms Brydges has her hands full *putting* on the finishing touches, […]. (CB)
d. He needed help *rethinking* his gendered, monolithic definition of family provider. (CB)
e. I wouldn't hesitate *moving* to Norway where he comes from. We're thinking of it and it's an exciting thought. (CB)
f. The only times to avoid travelling in the Metro are peak hours when you will see, and it is fascinating to watch, people performing the most extraordinary manoeuvres to get in (and out) of an already-full carriage; indeed, the guard will sometimes give the final coup de grâce with his boot to assist *closing* the doors. (CB)

g. A little tailor is sent on a quest with a glass key and succeeds *rescuing* a sleeping beauty, slaying the sorcerer and living happily ever after in the palace with his beautiful bride. (CB)

### *1.2.4 Dubious cases:* spend TIME + *IPC and related constructions*

Further, there are a number of construction types whose status as IPC-constructions is doubtful even among its most central members. The PCs in these constructions can be said to approach the adverbial end of the cline from IPC to adjunct. By far the most important group of constructions of this kind is clustered around the prototype illustrated in (39a), which has the verb *spend* in combination with a noun designating a period of time, and a PC designating the activity that the matrix clause subject is taken up in during that period of time. The status of the PCs in these constructions is dubious because the construction allows questioning by *how* (39b), but also wh-extraction (39c) and, arguably, questioning by *what* (39d). Insertion of an adverbial subordinator is not allowed (39e), but this may be due to the fact that English lacks a subordinator of manner. The construction is also remarkable, however, because omission of the PC often makes the sentence sound oddly incomplete (39f).

(39) a. He in turn would spend his time *boozing* and *nightclubbing* with mates […]. (CB)
b. How did he spend his time?
c. The mates he spent his time boozing and nightclubbing with.
d. (?) What did he spend his time in?
e. *He in turn would spend his time while/when boozing and nightclubbing with mates.
f. *He in turn would spend his time.

Variants of the construction arise when the time-NP is replaced by an NP denoting material goods, or when the verb *spend* alternates with the verbs *employ* (marginally), *lose*, *pass* and *waste*. Notice that with all of these variants omission of the PC is more acceptable. This indicates that the reason for the obligatoriness of the PC with *spend TIME* is pragmatic rather than syntactic: whereas all things existing spend time by definition and stating so is hardly informative, wasting time and spending goods are optional activities, alternative to using time and keeping goods, which suffices to make them of interest by themselves. Support for this view comes from the fact that *spend TIME* too can be used without an additional PC if the *TIME*-NP receives extra modification, as with the NP *evening* in (40).

(40) After the long conversation between herself and Lord Lackington which followed on the momentous confession of her identity, Julie spent a restless and weary evening, which passed into a restless and weary night. (1903, CEN)

Noteworthy is also that if the PC in a *spend*-construction is absent and the *TIME*-NP receives no extra modification, the PC's place will typically be taken by a PP or adverbial, as in (41). It is this paradigmatic relationship to clearly adverbial elements that suggests that the PCs in this constructions are (semi-)obligatory adverbials rather than IPCs proper.

(41) She arrived in Jamaica in April, intending to spend six months *there*. (CB)

That it is a thin line separating adverbials from IPCs, however, appears from the fact that some variants of the *spend TIME* construction do seem to take genuine IPCs. Compare in this respect examples (42a-b): example (42a) can still be interpreted as the negated answer to the question 'how did x waste time?'; in (42b), by contrast, the PC does not denote a manner of wasting time but an activity carried out by the subject without wasting time – that is, *waste no time* here functions as a verbal idiom specifying a relation of immediate (unhesitating) and intentional realisation between its subject and an action of which the subject is the agent. Accordingly, the PC in (42b) cannot be the focus of a *how*-question, and omission of the PC alters the meaning of the matrix predicate from 'not hesitate' to 'not idle'.

(42) a. Manchester United wasted no time *mourning* the loss of their Premiership crown. (CB)
b. Handball by a keeper outside his area in this competition warrants an instant dismissal and the referee wasted no time *waving* a red card. (CB)

### *1.2.5 More dubious cases*

Apart from the group of constructions clustered around the expression *spend TIME*, there are some other constructions whose PC allows wh-extraction but otherwise shows no signs of being a genuine IPC or whose IPC-status is at best questionable. Such constructions are therefore best classified as containing a participial adjunct. The underlying point here is that participial adjuncts do not readily combine with just any matrix clause. Instead, they tend to hold weak collocational ties with various more or less abstractly specified main clause con-

structions, in combination with which they form vaguely idiomatic expressions. Characteristic of the main clauses in question is that they contain only one participant (the subject) and therefore carry a relatively low information load. Examples are *go to bed* / *wake up* (43a), *die* / *get killed* (43b), *get* ADJ (43c), VERB of motion (43d), VERB of rest / motion + *around* (43e) and so on.

(43) a. You can go to bed *thinking* about something and wake up *thinking* about it next morning. (CB)
b. They're pulling statues down in Moscow these days. And to replace them there are men and women many Russians feel should be honored, such as Andrei Sakharov or the young men who died *resisting* last August's coup […]. (CB)
c. You may not get rich *backing* them, but put your cash elsewhere and you're throwing it away. (CB)
d. The woman enters *holding* a box of stockings. (CB)
e. At night workers just sat around *playing* cards or *sleeping*. (CB)

### *1.2.6 Conclusions*

In view of the preceding discussion, IPCs may be thought of as a gradient category, with some constructions exhibiting all defining features, others exhibiting some, and yet others none at all (cp. Quirk 2004 [1965]). How such a gradient might have to be conceived of is illustrated in Table 2, where a number of constructions are judged with respect to the defining characteristics of IPCs (versus adjuncts) and a few further characteristics that have surfaced in the preceding discussion.[8] Real IPC-constructions are found at the left hand side of the table, whereas constructions whose PC shows more adjunct-like behaviour are found at the right hand side.

There is a dangerous elegance to this kind of representation and the implicational hierarchies it might suggest. The gradient ordering in Table 2 is in some respects imperfect and furthermore glosses over certain syntactic differences between constructions whose occurrence the present gradient cannot predict. Thus there is the semi-obligatory nature of PCs with *spend TIME* that cuts across the IPC-gradient at an apparently arbitrary point. Then there is the occasional possibility of replacing a PC with a functionally equivalent PP or adverb such as *there*, again found with *spend TIME*, but also with *waste GOODS* and *be*

[8] Notice that in Table 2 characteristics a. (omission of PC renders matrix clause ungrammatical) and c. (omission of PC changes semantics of matrix clause predicate) are logically linked. If under omission of the PC the meaning of a predicate changes it still has to be grammatical, while if a predicate becomes ungrammatical, presumably we can no longer say anything about its meaning.

*better off* and, marginally, with *have no business* (though not *have no right*!). Similarly, looking at the possibility of questioning the PC by *what*, this operation seems fairly acceptable for *spend TIME* and *waste GOODS*, yet not for *be gone*, *be finished* and *have no business*, which are nevertheless higher up the IPC-cline in other respects – note though that because *what*-questioning requires the addition of a preposition to the matrix clause predicate, the outcome of the test depends on the availability of a prepositional alternative to the PC-construction and may reveal little about the PC as such. Finally, the gradient in Table 2 ignores the possibility that expressions like *be out* (like *be off* and *be gone*) approach yet another construction type – that of aspectual auxiliary constructions with verbs such as *lie* or *sit*.

**Table 2**: IPCs as a gradient category.

| *Enjoy* | *Have no business* | *Be tired* | *Be finished* | *Have trouble* | *Be busy* | *Be lucky* | *Be out* | *Spend TIME* | *Waste GOODS* | *Be better off* | *Make money* | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| + | ? | - | - | - | - | - | - | ? | - | - | - | i. Omission of PC renders matrix clause ungrammatical |
| + | + | + | ? | - | - | - | - | - | - | - | - | ii. No insertion of adverbial material between predicate and PC |
| n/a | ? | + | ? | + | ? | ? | ? | - | - | - | - | iii. Omission of PC changes the semantics of the matrix clause predicate |
| + | + | + | + | + | + | + | + | - | - | ? | - | iv. PC cannot be questioned by adverbial interrogative pronouns |
| + | + | + | + | + | + | + | + | + | ? | - | - | v. PC resists insertion of subordinating conjunctions |
| + | + | + | + | + | + | + | + | + | + | + | ? | vi. PC allows wh-extraction |

Be that as it may, the gradient in Table 2 certainly licenses the following conclusions: On the one hand, it is difficult (and perhaps impossible) to find criteria that give a unified description of IPCs as a category and strictly delimit them from other categories. On the other, there is no doubt that the PCs in certain environments show complement-like behaviour and differ from adverbial clauses, regardless of whether or not such PCs form a single coherent (if gradient) class. In that respect at least, an analysis of certain constructions in terms of IPCs is justified. A tentative suggestion as to how to interpret the gradient in Table 2 more meaningfully will be offered below (see section 3).

## 2 Sources and diffusion

### *2.1 Gradual emergence of IPCs*

IPCs are a relatively recent phenomenon in English and if they form a category at all, it is an emergent one. Two of the constructions discussed or mentioned in the preceding section go back a long time, but both are dubious as to whether they are in fact IPCs. Specifically, *done* is attested with *ing*-clauses as early as the sixteenth century, but only in the pattern where *done* is preceded by the auxiliary *have*, and it is not inconceivable that at this point the construction is in fact gerundial. A very early but hard-to-interpret example is given in (44a), a clearer instance of slightly later date is given in (44b). Evidence for the possible gerundial origin of the construction is given in (44c-d). *Ing*-clauses following the expression *be gone* turn up around the same time, but only in the fixed phrases *be gone a-hunting* / *a-fishing* / *a-birding*, as in (45a-b), with a characteristic *a-* prefix that renders the syntactic status of these forms dubious as well – moreover, as pointed out above, constructions of this kind may have come to lean closer to catenative uses of verbs of motion and rest than to IPCs proper, at least as far as Present-Day English is concerned.

(44) a. and when they had don plahyng, and then begane the sagbottes plahyng, (1553-59, PPCEME1)
b. 1591 Shakesp., Gent. II, iii, I, ‘twill be this hour ere I haue done *weeping* (quoted from Visser 1963-73: 2209)
c. soþ it is þat dymes ben due vn-to prestis in þe olde lawe, but þey weren holdun to do a3en *sleyng* of beestis & hard seruyss. (ICAMET)
‘it is true that tenths were due to priests according to the old law, but they were obliged to do their own slaughtering of animals and hard work.’
d. than risith on of the wisist lordis and reportith to the peple gret recomendacioun and preysyng of the kyng, and of þe good governaunce,

and done gret *thankyng* vnto god þat hath sent so excellent a witt vnto the kyng of Iewes to gouerne hem in suche wise (ICAMET)

(45) a. her husband is this morning gone a *Birding*: (1599, PPCEME2)
b. sure I thinke she be gone a *fishing* for her. (1630, PPCEME2)

Focusing on the most common IPC constructions as described in the previous section, the earliest straightforward examples attested date from the seventeenth century, when IPCs start to combine with the adjective *busy* (46a-b). The use of obligatory PCs with *spend TIME* (the IPC-status of which is again questionable, see above) shows up around the same time (46c), although it only becomes highly current in the course of the nineteenth century. Other types lag behind considerably, however: There are the IPCs combining with light verb idioms *have difficulty* or *have trouble* (46d), or IPCs with other adjectives than *busy* – including *done* (preceded by *be*) and *tired*, and possibly also *happy* and *late* (46e-f) – all of which first show up in the second half of the nineteenth century. The twentieth century sees further innovations, among which the use of IPCs with the adjectives *bored, comfortable, finished, fed up, slow* and *quick* and the light verb idioms *have success, have no right* and *have no business* (46g). Another clearly twentieth-century innovation is the occasional use of IPCs as complements or postmodifiers to nouns such as *difficulty, job, problem*, (*one's*) *time* or *work* (46h-i).

(46) a. [A]nd so home to supper -- my people busy *making* mince-pies (1666, PPCEME3)
b. Charles Smith and William Moon were both tried for stealing a Silver Tankard from one John Morris, value 6 l. they came to drink at Morris's House, and whilest the Man of the House was busy *waiting* on the other Guests, the Tankard was gone, and the Men too, without paying the Reckoning: […] (1693, *Proceedings of the Old Bailey*)
c. and she spent the whole day *making* herself clean. (1668, *Diary of Samuel Pepys*)
d. Dear Sir,--For a long time past I have had considerable difficulty *deciding* the important question, 'Who is the master of my own house? Myself, or YOUR SON Lupin?' (1894, CLMETEV3)
e. We were no sooner done *eating* than Cluny brought out an old, thumbed, greasy pack of cards, such as you may find in a mean inn, and his eyes brightened in his face as he proposed that we should fall to playing. (1886, CEN)
f. I am quite happy *standing* here alone in a crowd, knowing nobody! (1894, CLMETEV3)

g. "I was slow *figuring* it out," Leaphorn said. "I smelled something about Jackson. But I figured him to act like a Navajo and he was acting like a white man." (CB)
h. Work on the horses. Work around the yard. Work *trying* to get pregnant. And, nearly a full-time vocation in itself, work *raising* Marley. (*Marley and me*, p.81)
i. Terry […] said I would only ever get a job *collecting* supermarket trollies or *cleaning* donkey shit at an animal sanctuary (*The curious incident with the dog in the night time*, p.33)

On the face of it, the non-simultaneous appearance of IPCs with different types of heads is reminiscent of diffusional changes, in which a construction becomes increasingly frequent over time and is gradually matched to an increasing number of hosting constructions (cp. 'extension' in Harris & Campbell 1995: ch.5). In any case, as IPCs begin to occur in more and more environments, they run into other (and older) constructions that are equivalent or roughly similar in meaning and with which they have to compete. The diffusion of IPCs is especially marked by competition with gerund clauses introduced by the preposition *in*, which in many cases turn out to present a very adequate semantic and syntactic alternative to the IPC. Compare in this light the examples in (46) above with the following examples in (47):

(47) a. It was proved that the three Prisoners coming into the house of Temple, and calling for Wine, whilst Wilson and Pain were busie in *drinking*, Ellenor Davis makes use of the opportunity, taking the silver Salt-seller, marchs off unknown to her Companions, whereupon the said Temple missing his Salt seller, apprehends the said Wilson and Pain, as Accessary in the Theft; (1686, *Proceedings of the Old Bailey*)
b. A friend of his had spent much time in *composing* a book, and went to Sir Thomas to have his opinion of it; (1753, CLMETEV1)
c. I had more and more difficulty in *keeping* the fat landlady at arm's length, and the nasty child was well beaten one day for lingering about my door. (1888, CEN)
d. [O]ur cousin of France is happy in *having* a cavalier who is so fit to uphold his cause either with tongue or with sword. (1891, CEN)
e. Several agencies have complained that Turkey has been slow in *approving* projects. (CB)

**Table 3**: Percentage of IPCs to gerund clauses with *in*.

| Matrix predicate | 1850-1920 | 1990-1995 |
|---|---|---|
| *Be busy* | 93.9% | 100.0% |
| *Spend TIME* | 43.0% | 98.9% |
| *Be happy* | 25.0% | 98.9% |
| *Be late* | 24.2% | 46.9% |
| *Take turns* | 22.3% | 85.7% |
| *Have trouble* | 17.6% | 94.3% |
| *Have difficulty* | 1.5% | 72.1% |
| *Be engaged* | 1.3% | 2.8% |
| *Be slow* | 0.0% | 18.2% |
| *Be right* | 0.0% | 3.2% |
| *Be successful* | 0.0% | 0.0% |
| *Be justified* | 0.0% | 0.0% |

In order to give some idea of the competition between the two clause types, Table 3 documents the percentage of IPCs to gerund clauses introduced by *in* with the most common predicates taking either clause type. Figures are given for two historical periods, the transition from the nineteenth to the twentieth century (the period 1850-1922) and the end of the twentieth century (the period 1990-1995), on the basis of material from CLMETEV and CEN, and from CB respectively.[9]

[9] The manner of obtaining figures differs slightly for the two periods represented in Table 3, due to the different concordancing programs used to access different corpora: CEN and CLMETEV were accessed using Wordsmith Tools 3.0; CB could only be accessed using the interface that comes with the corpus. Differences apply to sampling methods and to the search strings used. Thus, the percentages in Table 3 are based on samples when corpus searches yielded too many instances. Samples were taken for *be busy* and *spend TIME* in the period 1850-1920, which were sampled at 1/3 hits and 1/12 hits respectively, and for *spend TIME*, *be happy*, *be late*, *be slow*, have *difficulty* and *have trouble* in the period 1990-1995, which were sampled at 200 hits for *be happy* and *be slow* and 150 hits for the other predicates (with the differences in sample sizes compensating for the amount of junk hits with some predicates). The search strings that were used were based on the main lexical word in the predicate (e.g. *busy*, *happy*, and so on) followed by a gap of zero to one words, followed by a word ending in *-ing* for the period 1850-1920, or a form tagged VBG (i.e. verbal form in *-ing*) for the period 1990-1995 – for the expression *spend TIME*, the gap between *spend*/*spends*/… and the form in *-ing* was zero to five words, so as to leave room for the *TIME*-NP. Note finally, that it is not always easy to distinguish between IPCs and participial adjuncts (see further section 2.2) or between the gerunds with *in* that function as complement and those that function as adverbial. For this reason, counts are based on instances that *can* be interpreted as IPCs or complement uses of the gerund with *in*.

Not all IPC constructions compete or compete exclusively with gerund clauses introduced by *in* (*tired* typically takes gerund clauses introduced by *of*; *happy* also takes gerund clauses introduced by *about*, *spend TIME* has also come to combine with gerunds introduced by *on*; and so forth). Moreover, the figures in Table 3 should be interpreted cautiously, as what is counted is potential rather than certain IPCs (i.e. constructions *allowing* an IPC-reading). This granted, Table 3 still provides a strong indication that at least with a number of predicates, the spread of IPCs progressed at the cost of the older gerund clauses with *in*. Most relevantly however, the table shows that the progression of IPCs did not proceed at the same time and at the same pace with all predicates. Clearly confirmed by the data in Table 3 is the finding that *busy* and – to a lesser degree – *spend TIME* started occurring much earlier with IPCs than other predicates. Another striking observation is that IPCs began to combine with some predicates such as *be engaged* already in the nineteenth century, but have failed to catch on in the course of the twentieth. For the twentieth century, too, it is intriguing to find that while some predicates have come to combine with IPCs exclusively, others are still highly resistant to use with IPCs, despite the fact that their use does not seem ungrammatical. For instance, constructions like those in (48), though apparently well-formed, remain unattested in CB:

(48) a. Values are important in career exploration because people who believe in the goals of their employers and whose values are in synch with their coworkers' values are more likely to be successful *getting* and *keeping* a job than those whose values conflict with others in the workplace. (Google)
b. One is justified *taking* the elevator one floor when no stairs are available. (Google)
c. This book was very helpful *getting* me up and running and doing what I wanted to do--make my website. (Google)

The questions that these findings raise are straightforward but remain hard to answer: First, what mechanisms give rise to the emergence of IPCs? And second, what determines the course and pace of diffusion? Put differently, why are some predicates affected by a change earlier and to a greater degree than others? While it is probably impossible to resolve these issues in all detail, some relevant insights can be gained by examining the mechanisms that could give rise to IPCs in particular environments, which is the main purpose of the remainder of this section.

### *2.2 The underlying mechanisms*

In all likelihood, the major mechanism in the emergence of the first IPC-constructions has been reanalysis from adverbial clauses. For the adjective *busy*, potentially ambiguous sequences are illustrated in (49).[10] If the presence or absence of a comma in writing is ignored, many instances of *busy* followed by a PC allow both an adverbial disjunct interpretation and an IPC-reading. On either reading the PC eventually describes what the subject of the matrix clause is busy doing, but on an IPC-reading, the PC restricts the semantic scope of the predicate by narrowing down the matrix clause subject's activity (as evoked by *busy*) to the activity denoted by the PC, while on a disjunct reading, the PC elaborates on the main clause and gives the additional information that justifies the speaker in calling the subject busy – information that typically consists in a description of the subject's current activity.

(49) a. Up, and to the office betimes; and there all the morning very busy, *causing* papers to be entered and sorted, to put the office in order against the Parliament. (1666, PPCEME3)

b. Thence took coach and I all alone to Hyde Park […], and so all the evening in the Park, being a little unwilling to be seen there, and at night home, and thereto W. Pen's and sat and talked there with his wife and children a good while, he being busy in his closet, I believe *preparing* his defence in Parliament, and so home to bed. (1668, *Diary of Samuel Pepys*)

c. When nature was most busie, the first weeke / *Swadling* the new-borne earth God seemd to like, / That she should sport her selfe sometimes, and play, / To mingle, and vary colours euery day. (a1631, John Donne, *An anatomy of the world*)

e. Mr. John Collins deposed, […] That he saw the Prisoner, who had a Black Patch upon his Nose, in the House Five or Six Minutes, very busy, *breaking* the Sashes and Frames of the Windows; (1716, *Proceedings of the Old Bailey*)

f. and by Course a new Consolidator being to be built, they were as busie as ever. *Bidding*, *Offering*, *Procuring*, *Buying*, *Selling*, and *Jobbing* of Feathers to who bid most; and notwithstanding several late wholesome and strict Laws against all manner of Collusion, Bribery and clandes-

[10] The construction with *busy* and IPC or disjunctive PC is particularly frequent in the diary of Samuel Pepys (49a-b), although as the examples show, it occurs elsewhere as well (49c-e). The same is true for the construction with *spend TIME*, which is also strikingly frequent in Samuel Pepys' diary; here too, however, the ambiguous instances are found both in the writings of Samuel Pepys (50a) and elsewhere (50b-c).

> tine Methods, in the Countries procuring these Feathers; never was the Moon in such an uproar about picking and culling the Feathers, such Bribery, such Drunkenness, such Caballing […], as the like has never been known. (1705, CEMET)

For *spend TIME* similar ambiguities are attested, originating from original disjunct uses, as is illustrated in (50). In each of the examples in (50), the PC may be read as a disjunct, elaborating on the matrix clause (whose pragmatic/syntactic requirement for some sort of adverbial is already met by another element), but may also be read as being itself part of the (semi-)obligatory adverbial modification of the *spend TIME* construction. Instances of this type could have given occasion for reanalysis as early as the seventeenth century – though of course, given that the PCs found with *spend TIME* are perhaps not strictly speaking IPCs, we may have to think of the proposed reanalysis as running from disjunct to obligatory adjunct, or we may even do without reanalysis altogether and limit ourselves to saying that adjuncts began to appear in this environment around the seventeenth century (but see below as to why this last possibility is somewhat less plausible).

(50) a. So to dinner to my Lady Sandwich's, and there after dinner above in the diningroom did spend an houre or two with her *talking* again about Creed's folly; (1665, *Diary of Samuel Pepys*)
b. Husbandry is another thing that doth occasion men to break; Some will spend their time in Drinking and Gaming, *neglecting* their business, until they are undone. (1681, LC)
c. but most of the little time I had with them was spent in a silent retiredness of spirit, *waiting* upon the Lord. (1683, CEMET)

For the adjectival predicates *tired* and *happy* a similar situation obtains, though here the ambiguity seems to lie between IPCs and adjuncts. Thus, PCs following *tired* (especially *get tired*), as in (51a-b), may be adjuncts specifying how or why the subject is getting tired, or IPCs specifying what the subject is getting tired of. Similarly, PCs following *happy*, as in (51c-d), may function as adjuncts describing the circumstances under which the subject is happy, or as IPCs denoting the source or object of the subject's happiness. Note that whereas *tired* is found in unambiguous IPC-constructions already in the nineteenth century (see examples (1b) and (32a) above), the evidence for *happy* is more difficult to interpret as the sequences with PC found in the nineteenth-century data are nearly all ambiguous between an adjunct and IPC reading, and convincing instances of the IPC construction are only available for Present-Day English (see (12b) and (30) above).

(51) a. But I think I'll try for the mule-buyer. I'm getting tired *looking* at these slab-sided cowmen. Now, just look at those mules--haven't had a harness on in a month. (1904, CEN)
b. I hope Hannah and John do not get *tired* doing my chores. (1903, CEN)
c. Yes, I see; but oh, I was so happy *being* a garden flower with the sunshine on my head, and I can't seem to care the least little bit for being a banian-tree! (1893, CEN)
d. Dearest mother, should we not be very happy *living* together in London? (1850, CLMETEV3)

Very similar ambiguities are found in the Present-Day English data for adjectives such as (*un*)*comfortable*, *bored* and *fed up*, and it is likely that these constructions went through the same development. Thus, the evidence in (52) for (*un*)*comfortable* is particularly suggestive: (52a-b) illustrate clear adjunct uses, (52c-d) represent the majority of present-day instances in allowing both an adjunct and IPC reading, and (52e-f) strongly favour the IPC reading. The variable interpretations compare neatly to those with the other emotive predicates: adjuncts specify the conditions or circumstances accompanying the situation of feeling comfortable as depicted by the main clause predication, while IPC constructions profile an attitudinal relation of willingness or reluctance between the main clause subject and the action denoted by the PC.

(52) a. [T]hough tall for a woman Jill was still short enough to be comfortable *sitting* on the bed. (CB)
b. Many immigrants couldn’t use traditional banks because they required a Social Security card, and they also felt more comfortable *doing* business in their own language. (CB)
c. it was clear that Deborah didn't feel comfortable *exploring* these issues at this juncture. (CB)
d. I think a lot of men would feel very uncomfortable *going* to work without a shirt and tie because in their particular field that's what gives them the confidence that they're part of of of the business world (CB)
e. I'm not sure we're looking yet at a real paradigm shift, but the accumulating evidence does make it very difficult to evade the conclusion that here, as in many other areas, there is a good deal more built in than most developmental psychologists had supposed (or felt comfortable *assuming*) a decade or two ago. (CB)

f. The psychiatrist impressed us as a sensitive and cautious man. After four visits, during which he played with Ted or interviewed Sara and me, he confessed that he was uncomfortable *making* a diagnosis. (CB)

For the adjective *late*, evidence is scantier, but it seems plausible that examples like (53) could give rise to reanalysis from adjunct to IPC. On the adjunct reading, the PC specifies an activity carried out at a late time – i.e. an activity concomitant to the subject's being late – while on an IPC reading, the PC specifies a goal that is associated with some (implicit) predefined time of realisation, and that the subject fails to achieve at that time (for a clear IPC instance, see (31c) above).

(53) a. Heavens, how sleepy I am! No wonder either! Late *going* to bed last night and up so early this morning. (1913, CEN)
b. He was merely a young man who had been rather late *visiting* one of the girls. (1893, CEN)

Finally, the light verb idioms *have difficulty* and *have trouble* show ambiguity along the same lines (although here too the evidence is less abundant), with PCs that can either be taken to describe the activity concomitant to the subject's having trouble (adjunct), or the activity the subject's trouble specifically pertains to (IPC), as is illustrated in the ambiguous examples in (54a-b). In such cases, adverbial PCs and IPCs are in practice indistinguishable. As (54c) conveniently illustrates, the adjunct reading of PCs is still available in Present-Day English, even though many instances will allow no adverbial interpretation at all (54d).

(54) a. Indeed, the captain told me he met difficulty enough *navigating* the shallow Main, and I think he prefers the deeper Rhine. (1910, CEN)
b. I will not tell the needless trouble I had *breaking* into that house--afterwards I found the front door was on the latch--nor how I ransacked every room for food, until just on the verge of despair, in what seemed to me to be a servant's bedroom, I found a rat-gnawed crust and two tins of pineapple. (1897, CLMETEV3)
c. [C]ustomers returning from Indonesia have experienced difficulties when *trying* to cash American Express traveller's cheques (CB)
d. I have difficulties *getting* marmalade in many hotels (CB)

As far as nineteenth century English is concerned, the only important exception is presented by *be done*, which acquires IPCs without any evidence of a preceding use of participial adjuncts with the same predicate. There is an obvi-

ous alternative explanation, however, since the already existent pattern *have done* with IPC (or gerund?) was in all likelihood extended to *be done* through analogy. In the nineteenth century *have done* occurs with and without IPC, as illustrated in (55a-b). On this basis, the use of *be done* without IPC in the same meaning as in (55c), attested from the end of the eighteenth century (Visser 1963-73: 2079), could be naturally extended to the use with IPC as in (55d).

(55) a. Interrupt me again, and I have done. (1884, CLMETEV3)
b. And when he has done *eating*, say I should like to have a few words with him, if he doesn't mind coming up here. (1873, CLMETEV3)
c. 1771 T. Jefferson, Letter to T. Adams in Harper's Magazine no. 483, One further favor and I am *done*. (quoted from Visser 1963: 2079)
d. We were no sooner done *eating* than Cluny brought out an old, thumbed, greasy pack of cards, such as you may find in a mean inn; and his eyes brightened in his face as he proposed that we should fall to playing. (1886, CEN)

Apart from *be done*, in all cases discussed, reanalysis gives a very natural account of the semantic and structural change at hand, and follows a recurrent pattern. On the one hand, the main clause predicates whose PC is reanalysed all implicitly evoke an elaboration site or e-site, i.e. a schematic participant that is activated along with a predicate (Langacker 1987, see also Keizer 2004). *Busy* implies an activity one is busy with; as emotive predicates, *happy* and *tired* imply a source that triggers the emotion they denote and in relation to which the emotion holds; the light verbs *have trouble* and *have difficulty* imply an intended goal on the part of their subject that circumstances hinder the subject from reaching, and so on. On the other hand, adverbial PCs with these predicates tend to map onto these e-sites by pragmatic implicature. Disjuncts used with *busy* are used to support the speaker's claim in the main clause by describing what the subject is busy with. Adjuncts with emotive predicates such as *happy* or *tired* strictly speaking only specify the circumstances under which the subject is *happy* or *tired*, but these often and naturally coincide with the source of happiness or tiredness. Adjuncts with *have trouble* or *have difficulty* describe an activity unfolding concomitantly to the subject's experiencing difficulty, and since such an unfolding activity is also related to the subject's intentions and goals – it is the activity the subject is trying to develop successfully – it naturally maps onto the intended goal implicitly evoked by the predicate. Eventually, what happens under reanalysis is that a pragmatic implicature becomes semanticised (cf. Traugott & König 1991; Lopez-Couso to appear). The primary consequence of this is that the mapping of the activity in the PC onto the e-site of the main clause becomes explicit and therefore restrictive. That is, the (schematic) activ-

ity filling the predicate's e-site now gets narrowed down to the activity in the PC. Roughly, the activity in the PC becomes the *only* activity that the subject is claimed to be busy with, happy or tired about, or experience trouble or difficulty with.

The invocation of an implicit e-site suggests that reanalysis may be guided by an underlying form of analogy (Hollmann 2003; Denison 2004; Fischer in prep.: ch.3, p.c.). After all, the e-site attributed to adjectives or light verb constructions in most cases already receives formal expression by other means than the IPC well before the reanalysis from adverbial to IPC takes place. For example, the use of IPCs with *be tired* virtually mirrors the older use of gerunds introduced by *of* with the same adjective. As pointed out above, nearly all IPC constructions have semantic equivalents of this kind in the form of prepositionally marked gerund constructions, and these equivalents may be thought of as providing the model on which reanalysis from adverbial to IPC takes place. This formulation cannot supplant an interpretation in terms of reanalysis: Firstly, if analogy were the only mechanism at work, it is unclear where IPCs could derive their form from – a clause with a verbal form in *-ing* occurring in a non-nominal position and unintroduced by any preposition: gerund clauses with a preposition could spread to new environments through analogy, but they could hardly drop their preposition through analogy. Secondly, an account based on analogy alone would have to start from the prepositional constructions that are the model of change, but if we look only at prepositional constructions, there is no recurrent pattern in the constructions that lose their preposition and those that do not. In such a view, the environments where IPCs turn up would seem to be selected randomly. The reanalysis account, by contrast, explains where IPCs get their formal characteristics from – viz. adverbial participle clauses – and why IPCs took on the distribution they did – viz. IPCs appeared where adverbials appeared. This being said, the reliance on an intervening analogical force that covertly guides reanalysis does not uselessly complicate matters, since it may indeed explain why reanalysis appears to take place so easily (i.e. on the basis of on the whole relatively infrequent combinations of matrix predicate and adverbial clause), and why reanalysis continually ends up creating competitors to already existing constructions. Other covert analogical influences may be posited as well. As analogy is in essence invisible, there is no telling to what extent the presence of one IPC construction in the grammar might have facilitated the subsequent emergence of other IPC constructions – or, indeed, to what extent the emergence of IPC constructions could have been helped along by the existence of other complement clauses based on an *-ing*-form of the verb such as gerund clauses.

Be that as it may, as represented here, the mechanism proposed to give rise to IPCs receives credibility from the fact that it accords well with general

observations from the history of English. That is, the emergence of IPCs instantiates a development towards closer integration in the matrix clause that is typical of adverbial clauses in general and that in each case appears to depend on the same interaction between implicit e-sites and the pragmatic implicatures conventionally attaching to adverbials in specific environments. For example, *to*-infinitives presumably derive from purposive adjuncts but have also acquired the possibility of functioning as verb complements (Los 2005), probably through reanalysis in environments such as (56a). A somewhat similar mechanism is likely to have given rise to the use of *lest*-clauses as complements (Lopez-Couso to appear), again with ambiguous environments serving as the trigger (56b). *For…to*-infinitives first showed up as complements with predicates ambiguous between an intransitive and a transitive reading, as in (56c) (De Smet to appear). Finally, the typological literature on complement clauses seems to confirm this pattern of development for other languages than English (Haspelmath 1989[11]; Croft 2001: ch.9). Obviously, the recurrent pattern of adverbials pragmatically mapping onto implicit e-sites and eventually turning into complements supports the plausibility of this scenario in the emergence of IPCs.

(56) a. & blodig regn & fyren fundiaþ þas eorþan *to forswylgenne & to forbærnenne*. <HomS 26 206>
‘and bloody rain and fire make haste / strive to devour and consume the earth.’ (quoted after Los 2005: 48)

b. Alyse me of Esaues handa, mines broðor, for þam ðe ic hyne swyðe ondræde, *þe læs ðe he cume & ofslea ðas modra mid heora cildum* (*DOEC*, c.1000, *Genesis* 32.11).
‘Deliver me from the hands of Esau, my brother, because I fear him very much, lest he come and kill the mothers with their children / because I fear him very much, (I fear) that he might come and kill the mothers with their children.’ (quoted after Lopez-Couso to appear)

c. First missis’s children fell ill of the measles, just when th’ week I’d asked for came, and I couldn’t leave them, for one and all cried *for me to nurse them*. (1848, E. Gaskell, *Mary Barton*) (De Smet to appear)

What can the account in terms of reanalysis tell us concerning the pattern of diffusion attested for IPCs? It is clear that the earliest instances of IPC-constructions mostly arose in environments that gave some occasion for reanalysis from adverbials to complement or complement-like clauses. It is therefore likely that reanalysis determined the initial distribution of IPCs over the set of

[11] Concerning Haspelmath’s universal pathway of change from purposive adverbial clause to infinitival complement it may be added that the change does not necessarily run its full course and may in fact recede again (Fischer 2000).

predicates potentially available for IPC-complementation. By spelling out its consequences, this conclusion can in turn serve to explain some facts about the order in which IPCs emerged in different environments, at the same time further corroborating the reanalysis hypothesis itself.

For example, the account in terms of reanalysis presented above rules out the initial occurrence of IPCs in at least one syntactic environment, well in accordance with the actually attested pattern of diffusion. If IPCs appeared in environments where they were reanalysed from participial adjuncts and disjuncts, it makes sense that they did not at first occur as noun complements or noun post-modifiers (as in Present-Day English examples (46h-i) above), since participial adjuncts or disjuncts do not attach to nominal heads. For the sake of illustration, consider the use of complements to the noun *difficulty* as illustrated in (57).

(57) a. Doctors often may not recognize the symptoms of clinical depression, which can be sadness, low energy, loss of interest in usual activities, difficulties *concentrating*, changes in eating or sleeping habits, and suicidal thoughts. (CB)
b. He accepted the defence's argument that Miss Short would not receive a fair trial because of the difficulties in *tracing* witnesses and evidence that would back her denial of the allegations. (CB)

Taking once more the percentage of IPCs (57a) to gerund clauses introduced by *in* (57b) as a measure of the success of IPC-constructions in a given lexico-grammatical environment, the lag of IPCs as noun complements is particularly striking: Table 4 compares the percentage of IPCs to gerund clauses with *in* as complement to the light verb idiom *have* / *find difficulty*(/*ies*) and as complement to the freely occurring nominal head *difficulty*(/*ies*), showing clearly that IPCs are far less advanced in the latter environment (see endnote 9 on counting practices).

**Table 4**: Percentage of IPCs to gerund clauses with *in*.

| Head | 1850-1920 | 1990-1995 |
|---|---|---|
| *Have difficulty* | 1.5% | 72.1% |
| *Difficulty* | 0.0% | 13.6% |

Diachronically, this indicates that IPCs are cascading down from the environments in which they were first reanalysed to new environments unaffected by the original reanalysis, with some form of analogical extension operating as the underlying mechanism. Thus, the initial emergence of IPCs through reanalysis

ties in with and partly explains the order of appearance of IPCs in different environments.

Reanalysis explains differences in timing in another respect as well. The account presented above distinguishes between disjunct-based reanalysis for *busy* and *spend TIME*, and adjunct-based reanalysis for *have difficulty*, *have trouble*, *be happy*, *be tired* and so on. In other words, IPCs with *busy* and IPCs or obligatory adjuncts with *spend TIME* are claimed to derive from disjuncts, while other IPCs are claimed to derive from adjuncts. This distinction is of some importance because it provides a possible explanation for the remarkable head start of *busy* in taking IPCs as compared to other IPC-taking predicates (cf. Table 3 and the examples in (46) above). Specifically, adjunctively used PCs are themselves a spreading construction, having gradually grown more frequent over the past four centuries. To show the progression of the construction, Table 5 gives estimates of the relative frequency of participial adjuncts in three subperiods.[12, 13] The possibility of reanalysis from adjuncts in a given environment is of course dependent on the recurrent presence of adjuncts in that environment. Although it is impossible and probably even nonsensical to formulate a critical frequency value above which reanalysis can start taking place, it may be rea-

[12] The figures for the period 1640-1710 are based on a subpart of PPCEME3, containing about 95,000 words of running text, on the basis of a corpus search on the orthographical sequence *-ing*. The figures for the period 1850-1920 were obtained in similar fashion, carrying out a similar search on *-ing* on a subpart of CLMETEV3 containing about 89,000 words. The figures for Present-Day English were obtained from the tagged and parsed ICE-GB corpus. The ICE-GB corpus was not searched for *-ing* (which the corpus interface does not allow) but for all verb phrases tagged as *-ing*-participles (ICE-GB was preferred to CB because the latter's interface neither allows searching on parts of words nor on tags unaccompanied by lexical material). To increase comparability with the historical data only subsections of the corpus were sampled, using different sampling rates to adjust the balance between text genres: the sections Non-academic writing, Reportage, Instructional writing and Persuasive writing were sampled at 25%, the section Creative writing was sampled at 50%. Together, the sampled sections are good for an estimated 69,000 words of text.

[13] Figures should be interpreted cautiously, since it is sometimes difficult to distinguish participial adjuncts from participial disjuncts. The presence of an explicit subject has been regarded as a fully reliable criterion to separate disjuncts from adjuncts. Less reliable is the presence or absence of a (written) comma; instead, in case of doubt the scopal tests described in section 1.1.1 were given priority in determining the adjunct/disjuncts status of a given adverbial participle. The estimated frequencies given in Table 5 are conservative, in that adjuncts have only been recognised as such when an adjunct reading yields the most plausible interpretation in context. One specific difficulty is presented by the Present-Day English data, which contain the fairly frequent use of the participle *using* in a somewhat grammaticalised form, as a semi-preposition roughly meaning 'with, by means of' (e.g. *do not open the can using a knife*). Because it is practically impossible to distinguish prepositional from non-prepositional uses, instances of this kind have not been counted. When included, the estimated frequency of participial adjuncts would rise to about 29 instances per 100,000 words for the period.

sonably assumed that the increase in the use of participial adjuncts, by introducing participial adjuncts into more environments, created more opportunities for reanalysis, and therefore bears indirect responsibility for the fairly simultaneous appearance of a variety of IPC-constructions around the end of the nineteenth century.

**Table 5**: Adjunctively used PCs (frequencies per 100,000 words).

| | 1640-1710 | 1850-1920 | 1990-1993 |
|---|---|---|---|
| Adjunct PCs | 3.3 | 13.5 | 20.2 |

Turning from adjuncts to disjuncts, and comparing the figures in Table 5 to the situation for disjunctively used PCs, it is found that the latter also saw an increase in use, but the main surge in frequency occurred much earlier in the transition from the Middle to the Early Modern period. Moreover, in all periods disjunctively used PCs vastly outnumbered their adjunctive counterparts (relative frequencies for disjuncts are found in Killie & Swan 2006). There is considerably less ground, therefore, to suspect that disjunct-based reanalysis would have been delayed by frequency-dependence in the same way as adjunct-based reanalysis. In combination, then, the differently timed developments of participial disjuncts and adjuncts, and the different starting points for reanalysis to IPCs can account for the time-lag between the use of IPCs with *busy* and other predicates. Furthermore, disjunct-based reanalysis may also account for the use of obligatory adjuncts with *spend TIME* at a time when adjuncts were still highly infrequent in other environments.

Other details of the order in which various predicates began to select IPCs can be explained in a similar fashion. For example, it has been suggested above that among the emotive predicates, *be tired* was probably the first to acquire IPCs, followed by *be happy* and other adjectival predicates. This sequence of events ties in with changes in the use of participial adjuncts, which not only grow in number in the course of the Late Modern period (as shown in Table 5 above), but also spread to new sub-uses. Already in the nineteenth century adjuncts quite readily combine with change of state predicates, especially with the verb *get*, specifying the manner or cause of the subject's moving from one state to another, as shown in (58). The use of adjuncts with (*get*) *tired* that probably gave rise to IPCs with the same adjective can be seen to instantiate this schematic pattern of co-occurrence (as is illustrated in (51a-b) above).

(58) a. You couldn't walk to-morrow if you took all the free samples of solid gold the boys would offer you. You'd get dizzy *looking* down prospect holes. (1905, CEN)
b. then Snowdon and a son as he had both got drowned *going* over a river at night. (1889, CEN)
c. I began to hope, although I'd got wrinkles *crying* about him. (1900, CEN)

By contrast, in the nineteenth century, predicates expressing an emotive state (rather than a change of state) only sporadically combine with adjuncts specifying the circumstances related to that state, as in (59a-b). This schematic pattern accounts for the adjuncts getting reanalysed to IPCs outside the change of state pattern, and the fact that it only became current in the course of the twentieth century correlates with the time of emergence of IPCs with emotive adjectives such as *happy* and *comfortable*.

(59) a. He could not go on with his honeymoon, so he would go up to London and work--he felt too miserable *hanging* about. (1910, CLMETEV3)
b. She remembered how she was reproved for peeping over her neighbour's shoulder, and how proud she felt *sitting* among all the workwomen. (1885, CEN)

Again, then, the dependence of the first IPC-constructions on reanalysis can be brought to explain facts about the specific order of emergence of different uses.

Despite the strong case that can be made for reanalysis, not all is reanalysis. In particular, analogical extension has played its part as well in the diffusion of IPCs. While some form of covert analogy may have supported reanalysis (see section 2.2 above), analogy is also likely to have operated as an independent mechanism of change and as such to have played a role in the development of IPCs outside environments licensing reanalysis. Thus, IPCs also appeared in a number of environments where their use is not foreshadowed by adverbial participle clauses. Assessing the influence of analogy, analogical extension seems to be a weak (or at least slow) force that mostly works locally (cp. De Smet to appear on *for…to*-infinitives), and that, in the case of IPCs, seems not yet to have run its full potential course. This is understandable since analogical extension has had relatively little time to regularise the use of IPCs across new environments since their first emergence. Still, while its effects are limited, analogy can probably not be ignored as an independent mechanism of change.

There are some relatively successful IPC-uses whose existence is most readily explained by analogical extension. There is the example of *be done* discussed earlier (see (55) above), whose use with IPCs from the end of the nine-

teenth century onwards is probably not the result of reanalysis but is based on the use of IPCs (or some kind of *-ing*-clause) with *have done*, and which in turn is likely to have formed the basis for the occasional use of IPCs with *be finished* in Present-Day English. Another example may be the use of IPCs with *have no right* and *have no business*, which could analogically derive from the use of IPCs with other light verb constructions such as *have* (*no*) *trouble* / *success* etc. The analogical link is not altogether straightforward, but syntactic farfetchedness may here be iconic of the strong emotional load these expressions tend to convey (*have no right* and *have no business* typically signal indignation) – that is, marked syntax is motivated by marked semantics.

In other cases, likely instances of analogical extension are somewhat less successful, so that while analogy can be imagined to foster all kinds of new IPC uses, in practice it mostly appears to remain somewhat ineffective. A truly innovative but otherwise relatively unsuccessful case of analogical extension is presented by the use of IPCs with simplex verbs, such as *assist*, *hesitate* or *succeed*, as in (60) (or (38e-g) above), the occurrence of which has so far remained restricted to a few scattered instances in the corpus material. Note further that, as with *have no right* and *have no business*, the analogical models on which the simplex verb formations are based are not altogether straightforward, although there do exist semantic and formal relations linking the simplex verbs to other, more common IPC-taking predicates. Semantically, *assist* can be linked to *need help* (see example above), *hesitate* to *have no hesitation* which in turn relates naturally to *have* (*no*) *difficulty*, *succeed* to *have* (*no*) *success* which in turn is again linked to *have* (*no*) *difficulty*. Formally, *assist* and *succeed* (though not *hesitate*) can be construed with gerund clauses introduced by the preposition *in*, just like many other IPC-taking predicates.

(60) The Town Council, who are the Local Health /Authorities, have been condemning many old properties recently, and are pressing proprietors to put their houses in order; but they have been considerably hampered by the great scarcity of houses, and naturally hesitate *taking* action which would have people removed with no place to house them. (Google)

Another likely instance of analogical extension is presented by the use of IPCs with *engaged*, *occupied*, and *employed* as in (61), probably based on the use of IPCs with *busy*. But as the figures in Table 3 above indicate, here too analogical pressure seems to be insufficient to forward these uses beyond occasional occurrence.

(61) a. Winter in a Flat racing yard is the most hated time of year for stable lads. It's the time they're engaged *breaking* in the yearlings and, apart from the odd all-weather card, there are no race meetings to break the monotony. (CB)
b. […] in a densely peopled quarter of the city, such as in our own day we should call a slum, where folk were employed *making* those articles which ministered to the comfort of the luxury of the more fortunate […]. (1903, CEN)
c. He had hurled his lasso with the rest, and it was trailing. He jerked about and fled for a mile or more, holding on with his legs while both hands were occupied *gathering* in the rope and *coiling* it about the high pommel of his saddle. (1898, CEN)

Similarly, analogical extension is likely to have been at play in the use of IPCs as noun complement to, for instance, the noun *difficulty*. Again, however, the data discussed earlier (see in particular Table 4 above) reveals that the extension of IPCs from one environment to another is hesitant. This lack of productivity or 'extendability' is most conspicuous, finally, for emotive adjectives: while these seem to be well-represented among the IPC-taking predicates and seem to form a natural basis for analogical extension, there still exist a good number of common emotive adjectives that remain unattested in the pattern in CB, including *annoyed*, *delighted*, *excited*, *glad*, *sad*, *sorry*, *surprised*, *upset* and so on. In sum, a number of innovative IPC uses are most easily explained as the outcome of a process of analogical extension, and these extensions seem to assert IPCs as a (mildly) productive syntactic pattern. At the same time, it is certainly not the case that analogy causes IPCs to disperse easily or quickly from one lexico-grammatical environment to another.

## 3 What are IPCs?

Having given a synchronic characterisation of IPCs and having examined their origins, it is time to tie together the previous findings and come to a view of what exactly IPCs are. Based on the preceding discussion, the following two-part definition can be given:

(i) a. Synchronically, IPCs are non-finite clauses with a main verb in *-ing* that differ from adverbial participial clauses in showing complement-like behaviour, but differ from most other complement types in not showing any nominal features.
b. Diachronically, IPCs derive from adverbial participial clauses that became interpretationally dependent on the semantics of their matrix

> clause predicates through reanalysis. After reanalysis had given rise to the first IPC constructions, analogy contributed to further spreading IPCs to a number of new environments.

While this minimal definition more or less delimits IPCs for analytical purposes, it says nothing about the ontological status of IPCs. In other words, does the category of IPCs correspond to anything that might be considered linguistically (if perhaps not philosophically) real?

### *3.1 Diachronic status*

Let us begin by looking at the diachronic side of the issue. As a diachronic phenomenon, IPCs can be situated at the cross-roads between a number of general characteristics of language and the specific history of English. Following Croft (2001: ch.9), language users have different ways of conceiving of (or 'construing') the relation between two situations, which is grammatically reflected in, among other things, the difference between adverbials and complements: in a complement construction one situation is perceived as a core participant in another situation (e-site elaboration), while in an adverbial construction one situation functions as the background against which the other situation unfolds (figure-ground configuration). Although adverbial and complement clauses are in this view conceptually distinct, there are at least two reasons why the boundaries between the two are not too strict. On the one hand, e-site elaboration is not an all-or-nothing affair. Because the requirement of a complement is dependent on the expectation of an extra participant linked to the complement-taking head, not all complements are 'obligatory' to the same degree (Langacker 1987; Keizer 2004) and e-sites often remain unelaborated. On the other, it has been observed that "adjunct-less intransitive clauses are rare" (Quirk *et al.* 1985: 506) – in other words, intransitive predicates are unlikely to occur without any further adverbial modification. In combination, implicit e-sites and the relative obligatoriness of adverbials form a breeding-ground for the emergence of new complement-types – a development further fuelled by the existence of other complement constructions that may serve as (covert) analogical models. As far as English is concerned, it has been argued here that the rise in discourse frequency of new adverbial types – participial disjuncts from the beginning of the Early Modern period, as well as participial adjuncts from about the nineteenth century – boiled over into the appearance of a new complement type – IPCs. Diachronically, then, the status of IPCs is that of an instance of a natural historical development.

### *3.2 Synchronic status*

The synchronic status of IPCs is more difficult to assess, as IPCs confront the analyst with synchronic vagueness that extends in at least two directions. On one, 'horizontal', axis, IPCs have been shown to take variable positions along a gradient from adverbial to complement (see Table 2 above). The question to be asked here is what is the meaning of this gradient? On another, 'vertical', axis, we need to find out whether IPCs as such form a linguistic category that is representative of speakers' internalised knowledge of language. In other words, do IPCs function as a language-specific construction?

In order to deal with the gradient along which IPCs are situated, a plausible solution is to relate the ordering of various predicates to the conceptual cline running from typical figure-ground configurations to e-site elaboration. It is a matter of degree whether a situation is seen as providing the background against which another situation is profiled, or as one element in an interaction with another participant. However, while it is easy to see why the conceptual boundaries between these two relation types are non-discrete, there is no easy way to account (in a non-circular way) for the specific positions that predicates take along the cline from one relation type to the other. The factor most likely to matter is of course the semantics of the IPC-taking predicate involved. That is, somehow the semantics of different IPC-taking predicates lead to those predicates assuming different positions on the gradient from adverbial to complement.

Importantly, this view is independently supported by the diachronic behaviour of the adverbial-to-complement gradient. If the position of a predicate on the gradient depends on its lexical semantics, we do not expect predicates to move along the gradient unless their meaning changes. Indeed, there is no indication that the IPCs with *busy*, once reanalysed as such, went on to become more resistant to omission in the course of the last four centuries; that the obligatory adjuncts with *spend TIME* gradually became more resistant to questioning by an adverbial interrogative; and so on. That is, once reanalysed, IPC-constructions stay where they are, fixed in some stable position along the adverbial-to-IPC cline. The fact that different constructions seem to be more or less ideal members of the class of IPC constructions, then, is not due to a diachronic development causing various elements to gradually move along a cline toward greater IPC-construction-hood. Rather, it is due to the lexical properties or requirements of the specific matrix predicates involved. The occasional exception proves the rule. As hinted earlier, some instances of the *spend TIME* construction have in fact moved up the cline from adverbial to complement, but this change is accompanied by semantic change in the matrix clause predicate itself – see the discussion of the idiomatised phrase *waste no time*, meaning 'not to hesitate' (section 1.2.4, contrasting examples (42a-b)).

But how precisely can the semantics of IPC-taking predicates be linked to a cline from figure-ground construal to e-site elaboration? It is not unthinkable that in such a 'lexical' matter no generalisation is possible. Yet on a tentative note, I would suggest that what ultimately determines the position of a predicate is the strength or impact of the interaction it expresses between two participants – that is, the degree to which the state of the subject can be conceived to be dependent on or responsible for the state of the situation expressed in the PC. The more clearly the process denoted by a predicate follows from a specified state in one participant and results in a specified state in the other participant, the higher the probability that the predicate will indeed be perceived as designating an interaction between the two participants, and the stronger and more intrinsic the cognitive link that will activate (the expectation of) the second participant essential to a relationship of e-site elaboration. The hierarchy suggested resembles Givón's (1980) binding scale or Hopper and Thompson's (1980) scale of transitivity, but only in part since it does not immediately matter whether influence runs from the first participant to the second or vice versa. Primarily, what matters is whether and to what degree some influence or interaction is likely to be perceived by the language user. For example, *be busy* and *have trouble* are ranked higher than *spend TIME* because they imply greater activity and intentionality on the part of the subject and thereby create a stronger link to a second participant in the form of a particular goal. In turn *be tired* and other emotive predicates are ranked slightly higher than *be busy* and *have trouble* because while the outcome of the subject's goal-oriented activity is indeterminate in the case of *be busy* and *have trouble*, there is an outspoken effect of one participant (the situation in the PC) on the other (the subject) in the case of *be tired*: the PC designates a source that triggers a certain emotion in the subject.[14] By hypothesis, a scale of this kind from weaker to stronger interaction may provide the required semantic underpinning of the gradient attested in the data.

Assuming, then, that the prototypical organisation of categories sufficiently excuses the somewhat disparate behaviour of IPCs, we may turn to the second question and ask whether IPCs form a synchronic category. For this to be the case, IPCs would have to meet the requirement that they all instantiate some more abstract pairing of form and meaning (Langacker 1987; Goldberg 2006). Such categorial status could be attributed to IPCs if they show some distinctive formal feature, a semantic idiosyncrasy unpredictable from their formal make-up (in the sense of Goldberg 1995), or some distributional restrictions that could be specified coherently (in the sense of Pawley and Syder 1983). If recognised as a category, IPCs could receive their own place among other complement types, which can typically be shown to carry their own meanings and be stratified over

[14] Incidentally, as Langacker (1991b: 326-7) points out, an experiencer is to some extant an active participant in event structure since experiencing requires the establishment of mental contact with the thing experienced as well as the activation of some cognitive representation (cp. Hollmann 2003: 59).

the inventory of complement-taking predicates accordingly (see e.g. Bolinger 1968; Givón 1980; Noonan 1985). Consequently, a number of views are logically possible:

(ii) a. IPCs fall apart as a synchronically unrelated set of purely idiomatic complement constructions with limited productivity
b. IPCs are an independent and internally coherent category, a distinct complement type in the system of complementation
c. IPCs are marginal members of the larger category of *-ing*-complements

Adopting a view of language as organised along constructional hierarchies, there is even a fourth option stating that the three possibilities under (ii) are simultaneously true albeit to various degrees. Although, strictly speaking, the issue cannot be solved conclusively on the basis of the available data, it seems that the last option is in fact the most plausible, with alternatives (ii.a) and (ii.c) asserting themselves most forcefully.

Relying on distributional evidence, it is not difficult to see IPCs as occurring in a set of separate semi-productive idioms: clearly a number of different prominent schemas exist around which IPC-taking constructions are clustered, most prominently the adjectival predicates (*be busy*, *be tired* and so on), the light verbs (*have difficulty*, *have trouble* and so on), and – if accepted as genuine IPC-taking predicates – the *spend* verbs (see 1.2 above). This ignores the existence of various other constructions taking IPCs, but even if the latter are included in the picture, the strong association of IPCs to a number of half-schematic environments is evident. To the extent that these environments are unrelated to one another, IPCs are heterogeneous in distributional terms.

Semantic arguments can be used to support an entirely different position. Some IPC uses can be identified as idiomatic on semantic grounds because their component parts do not predict the meaning of the composite construction, e.g. the IPC constructions with *have no business* or *waste no time*. This granted, however, it is generally true that IPC constructions do not carry any semantic specifications beyond what is predictable from their component parts. None of the matrix clause predicates combining with IPCs have to be attributed senses they do not have in other constructions. Furthermore, IPCs themselves appear not to differ semantically from other *-ing*-clauses. Attempts to capture the meaning of the *-ing*-suffix can be safely extended to include IPCs (e.g. Langacker's semantic characterisation of the suffix as construing the process denoted by the verb stem "holistically" and, at the same time, "as being effectively homogeneous" as "a representative series of internal states" (1991: 443)).[15] In semantic

[15] Admittedly, as far as IPCs are concerned such characterisations remain hard to prove on non-intuitive grounds. Evidence for the semantic characterisation of the *-ing*-suffix comes

terms, then, IPC constructions do not differ in any obvious way from other patterns of *-ing*-complementation.

As for the formal arguments, IPCs are distinguished from nearly all other complement clauses by the *-ing*-form of their main verb and from gerund clauses by their non-nominal character. This looks like a strong argument in favour of viewing IPCs as an independent category, but notice that the formal distinction between IPCs and gerunds is in fact based on covert criteria, and therefore less persuasive than the distinction from other complement types based on the overt presence of *-ing*. Covert criteria may in fact be ignored when doing so results in a unified analysis of otherwise identical patterns (cp. Goldberg 2006: ch.2). Moreover, the nominal nature of the gerund is itself not above suspicion. As is well known, gerunds have been losing nominal features throughout their history. This is true not only of the internal syntax of gerunds (Fanego 2004), but also of their distribution over syntactic environments. Thus, over the past centuries, evidently un-nominal uses have surfaced that would be poorly analysed if treated separately from more clearly nominal examples. For instance, gerunds have come to be used in extraposition and pseudo-extraposition constructions, as in (62a), have spread to various idiomatic environments where they no longer have a strictly nominal alternative, as in (62b), and at least in Present-Day English even occasionally allow subject raising, as in (62c). None of the gerunds in (62) can be replaced by an ordinary NP.

(62) a. Alice laughed. ‘There’s no use *trying*,’ she said: ‘one CAN’T believe impossible things.’ (1865, CLMETEV2)
b. Luckily, Boomerangs have a habit of *returning*. (CB)
c. Nearly two tons of pure cocaine has been stopped *getting* on the market. (CB)

The extent to which IPCs are seen as a formally distinct category, then, depends on the importance attached to covert syntactic criteria. Even if covert criteria are accepted, however, it is difficult to isolate IPCs entirely from gerund constructions since other *ing*-clauses whose origins are more clearly gerundial also show occasional non-nominal behaviour.

The synchronic ambivalence of IPCs is reflected diachronically in the mechanisms that gave rise to them. Being primarily dependent on reanalysis, the emergence of the first IPCs could theoretically have occurred in different environments independently. The prominence of reanalysis can therefore be invoked to support a view of IPCs as a loose set of separate constructions. However, the more recent instances of analogical extension are indicative of a certain degree

---

from the distribution of the progressive tense and from the mass noun-like behaviour of gerunds in a number of environments (Langacker 1991; Heyvaert 2004). It may be assumed that other *-ing*-constructions inherit their semantics from such uses.

of productivity beyond isolated, semi-idiomatic construction types, which supports a view of IPCs as an independent category, at least for Present-Day English. Moreover, since reanalysis may have partly depended on a form of hidden analogy anyway, there is no reason to exclude the possibility of a broad category – however marginal – that encompasses both nominal and non-nominal *-ing*-complements.

## 4 Conclusion

It has been suggested that grammars change because they leak (Denison 2004). True as this is, the opposite is also true: grammars leak because they change. IPCs instantiate a change from adverbial to complement. This change requires a certain amount of gradience in that it depends on the possibility of construing certain predicates either as transitive or intransitive. In the terminology of the preceding discussion, the change from adverbial to IPC is possible because predicates have an elaboration site that is sufficiently weak to be left implicit but sufficiently strong to assert itself and attract a new complement-type. At the same time, the change from adverbial to IPC creates gradience, in that a new set of constructions emerges whose categorial status confuses existing categorial boundaries – in particular, the boundaries of the English gerund as a nominalisation – and also creates new fuzzy boundaries – in particular, the boundaries between the IPC constructions themselves. Finally, the IPCs that get reanalysed fall by their nature into patterns of gradience related to the non-discrete relationship between complementation and adverbial modification.

Frustrating though it may be, the finding that IPCs refuse to form a neat category is not unlikely to reflect speakers' actual linguistic knowledge. Such knowledge, after all, is generalised from much the same data as the present analysis.

## 5 List of references

Bolinger, Dwight. 1968. Entailment and the meaning of structures. *Glossa* 2: 119-127.

Croft, William. 2001. *Radical construction grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Denison, David. 2004. Do grammars change when they leak? In Christian J. Kay, Simon Horobin and Jeremy Smith (eds). *New perspectives on English historical linguistics*. Amsterdam: John Benjamins. 15-29.

De Smet, Hendrik. To appear. *For…to*-infinitives as verbal complements in Late Modern and Present-Day English. Between motivation and change. *English Studies*.

Fanego, Teresa. 2004. On reanalysis and actualization in syntactic change. The rise and development of English verbal gerunds. *Diachronica* 21: 5-55.

Fischer, Olga. 2000. Grammaticalisation: unidirectional, non-reversable? The case of *to* before the infinitive in English. In Olga Fischer, Anette Rosenbach and Dieter Stein (eds). *Pathways of change. Grammaticalization in English*. Amsterdam: Benjamins. 149-169.

Fischer, Olga. In preparation. *Approaches to morphosyntactic change from a functional and formal perspective*. Oxford: Oxford University Press.

Givón, Talmy. 1980. The binding hierarchy and the typology of complements. *Studies in Language* 4: 333-377.

Goldberg, Adele E. 1995. *Constructions. A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. 2006. *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.

Greenbaum, Sidney. 1969. *Studies in English adverbial usage*. London: Longmans, Green and Co.

Haspelmath, Martin. 1989. From purposive to infinitive – a universal path of grammaticization. *Folia Linguistica Historica* 10: 287-310.

Heyvaert, Liesbet. 2004. Towards a symbolic typology of *-ing* nominalizations. Michel Achard and Suzanne Kemmer (eds). *Language, culture, and mind*. Stanford: CSLI Publications. 493-506.

Hollmann, Willem B. 2003. Synchrony and diachrony of English periphrastic causatives. A cognitive perspective. Doctoral thesis. University of Manchester: Faculty of Arts.

Hopper, Paul J. and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56: 251-299.

Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.

Keizer, Evelien. 2004. Postnominal PP complements and modifiers. A cognitive distinction. *English Language and Linguistics* 8: 323-350.

Killie, Kristin and Toril Swan. 2006. Pragmatics and the development of adverbial *-ing*-clauses (converbs) in English. Paper presented at ICAME 27, Helsinki, 26 May.

Kortmann, Bernd. 1991. *Free adjuncts and absolutes in English. Problems of control and interpretation*. London: Routledge.

Langacker, Ronald W. 1987. *Foundations of cognitive grammar*. Vol.1. *Theoretical prerequisites*. Stanford: Stanford University Press.

Langacker, Ronald W. 1991. *Foundations of cognitive grammar*. Vol.2. *Descriptive application*. Stanford: Stanford University Press.

Lopez-Couso, María José. To appear. Adverbial connectives within and beyond adverbial subordination. The history of *lest*. In Ursula Lenker and Anneli Meurman-Solin (eds). *Clausal connection in the history of English. Current issues in linguistic theory*. Amsterdam: John Benjamins.

Los, Bettelou. 2005. *The rise of the* to-*infinitive*. Oxford: Oxford University Press.

Noonan, Michael. 1985. Complementation. In Timothy Shopen (ed.). *Language typology and syntactic description*. Vol.2. *Complex constructions*. Cambridge: Cambridge University Press. 42-140.

Pawley, Andrew and Frances H. Syder. 1983. Two puzzles for linguistic theory. Nativelike selection and nativelike fluency. In Jack Richards and Richard Schmidt (eds). *Language and communication*. London: Longman. 191-226.

Quirk, Randolph. 2004 [1965]. Descriptive statement and serial relationship. In Bas Aarts, David Denison, Evelien Keizer and Gergana Popova (eds). *Fuzzy grammar*. *A reader*. Oxford: Oxford University Press. 327-349.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.

Ross, John Robert. 1973. Nouniness. In Osamu Fujimara (ed.). *Three dimensions of linguistic research*. Tokyo: TEC Company. 137-257.

Taylor, John R. 1998. Syntactic constructions as prototype categories. In Michael Tomasello (ed.). *The new psychology of language. Cognitive and functional approaches to language structure*. Vol. 1. London: Lawrence Erlbaum Associates. 177-202.

Traugott, Elizabeth C. and Ekkehard König. 1991. The semantics-pragmatics of grammaticalization revisited. In Elizabeth C. Traugott and Bernd Heine (eds). *Approaches to grammaticalization*. Vol.1. Amsterdam: John Benjamins. 189-218.

Visser, Fredericus Theodorus. 1963-73. *An historical syntax of the English language*. Leiden: Brill.

# Automatic measurement of speech rate in spoken Dutch*

**Nivja H. de Jong**
*University of Amsterdam*

**Ton Wempe**
*University of Amsterdam*

*In this paper, we describe a method to automatically measure speech rate without the need of a transcription. A script written in the software program PRAAT detects syllables in running speech. Peaks in intensity (dB) that are preceded by dips in intensity are considered as potential syllable nuclei. The script subsequently deletes peaks that are not voiced. Testing the resulting syllable counts of this script on two corpora of spoken Dutch, we obtained high correlations between speech rate calculated from human syllable counts and from automatically determined syllable counts. We conclude that a syllable count measured in this automatic fashion, suffices to reliably assess and compare speech rates between participants and tasks.*

## 1 Introduction

Becoming fluent in a second language is one of the most difficult aspects of learning a second language (J. H. A. L. De Jong & Van Ginkel, 1992). At the same time, measures of fluency are an important aspect of second language speaking proficiency (see, e.g., the Common European Framework of Reference for Languages (2001), p. 28 – 29). Tavakoli and Skehan (2005) have suggested that three different aspects of fluency can be distinguished: breakdown fluency (number and length of pauses), speed and density per time unit (speech rate), and repair fluency (false starts and repetitions).

For most speaking tests, fluency is a score awarded by human raters who presumably use all aspects of fluency in their judgement. Correlations between

* This research was carried out at the Amsterdam Center for Language and Communication. Research funded by NWO under grant number 575-21-009 awarded to Prof. J.H. Hulstijn and Dr. R. Schoonen ("Unraveling second language proficiency"). We wish to thank Renske Berns for her help in counting syllables in a subset of the WiSP-corpus, and Jan Hulstijn, Rob Schoonen, Rob van Son and Paul Boersma for their helpful comments on earlier versions of this paper.

such subjective measures of fluency and objective measures of fluency have shown that *speech rate* is the best predictor of subjective fluency (Cucchiarini, Strik, & Boves, 2002). Kormos and Dénes (2004) show that speech rate in terms of number of syllables per time unit is a good predictor of subjective fluency. In order to ensure objective scores on speaking tests, objective measures of fluency would be preferred. 'Breakdown fluency' can objectively be measured by measuring the duration and number of silences in running speech, 'repair fluency' can be determined objectively by counting false starts and repetitions, and speech rate, finally, can objectively be measured by counting syllables. However, counting syllables is a tedious job and is often cast aside due to time constraints. In the context of a large-scale research project on the correlates of speaking proficiency carried out at the University of Amsterdam (What is Speaking Proficiency: http://www.hum.uva.nl/wisp), we developed two tools to measure fluency automatically. We wrote a script in the software program PRAAT to automatically detect silence in speech (a simplified version of which is now incorporated in the button *To TextGrid (silences)*. This paper concerns another script in PRAAT that automatically detects syllable nuclei to compute speech rate in terms of syllables per time unit.

Besides for second language research that (wishes to) include a measure of fluency, speech rate is a very important signal for automatic speech recognition as well. Human listeners are able to understand both fast and slow speech in an automatic way. Speech recognizers implemented in computers, however, perform relatively poorly when speech rate is very fast or very slow. In order to improve computer performance, several researchers have proposed that measuring speech rate prior to speech recognition will result in higher success rates of automatic speech recognizers (Pfau, Faltlhauser, & Ruske, 2000) and several ways to automatically measure speech rate in terms of phones and/or syllables per time unit have been put forward.

Mermelstein (1975) developed an algorithm to segment speech into syllables by finding minima in loudness that serve as possible syllable boundaries. Verhasselt and Martens (1996) presented an automatic speech detector that measures phone boundaries and thus calculates rate of speech as phone rate. The phone boundaries are provided by a Multi Layer Perceptron that is trained on a subset of the data that must be hand-segmented at the phone level. Pfitzinger (1999) uses a combination of syllable rate and phone rate to correlate with perceptual speech rate. Syllable rate is calculated by counting peaks in the energy contour, while phone rate is calculated by use of transcription. The syllable, phone, and perceptual speech rates were measured over (very) short stimuli (625 ms). Hunt (1993) used recurrent neural networks to detect syllables. Pfau and Ruske (1998) determined syllable nuclei by detecting vowels on smoothed modified loudness and then calculated speech rate.

All of these different automatic ways to measure syllable and/or phone rate are quite successful. It is difficult, however, to compare the success of these automatic measurers, because they were all used on different corpora, and their success was reported in different ways. Some researchers report correlations between human and automatic speech rate, others report a percentage of syllables (or phones) undetected and falsely detected as compared to human measured syllables (or phones), and yet others report a correlation between the number of manually measured and automatically measured syllables (or phones). The difference in corpora used should also be noted, as some studies used large corpora with many different speakers and others used quite small corpora with few speakers; some used corpora of speech read aloud while others used (semi-) spontaneous-speech corpora. Finally, a noteworthy difference between studies concerns the length of the spurt on which speech rate was calculated. Some studies used extremely short time-windows to calculate speech rate, and others used much longer windows. Perhaps the most obvious reason we cannot compare success of these different automatic speech rate measurers is that the length of the time-window (or spurt) as well as the variance in spurt length will strongly influence calculations of speech rate.

Many of the proposed speech rate measurers need to be trained on a subset of the data that is transcribed or preprocessed by hand (Hunt, 1993; Pfau & Ruske, 1998; Pfitzinger, 1999; Verhasselt & Martens, 1996). In this paper, we will present an easy way to automatically measure speech rate without the use of preprocessing or the need for transcriptions and test it on two different corpora of spontaneously spoken Dutch. To be able to compare the success of the script over these two different corpora, spurt length was controlled. We wrote a script in PRAAT (Boersma & Weenink, 2007) using a combination of intensity (similar to Pfitzinger, 1999) and voicedness (similar to Pfau and Ruske, 1998) to find syllable nuclei.

## 2 The algorithm

In what follows, we describe the subsequent actions the script completes to find syllable nuclei using intensity (dB) and voicedness. Before the script is run, sound files that are quite noisy should be filtered so that the frequency range is speech-band limited.

Step 1. We extract the intensity, with the parameter ‘minimum Pitch’ set to 50 Hz and using autocorrelation, hence using a window size of 64 ms, using time steps of 16 ms.

Step 2. We consider all peaks above a certain threshold in intensity as potential syllables. We set the threshold to 0 or 2 dB above the median intensity measured over the total sound file (0 dB if the sound is not filtered, 2 dB if the sound is

filtered). We use the median, rather than the mean, to calculate the threshold in order to avoid including extreme peaks in the calculation of the threshold.
Step 3. We inspect the preceding dip in intensity and only consider a peak with a preceding dip of at least 2 or 4 dB with respect to the current peak as a potential syllable (2 dB if the sound is not filtered, 4 dB if the sound is filtered).
Step 4. We extract the Pitch contour, this time using a window size of 100 ms and 20 ms time steps and exclude all peaks that are unvoiced.
Step 5. The remaining peaks are considered syllable nuclei and are saved in a TextGrid (point tier).

The script is available on the personal webpage of the first author (De Jong & Wempe, 2007).

Figure 1 shows a part of a sound file together with the output as TextGrid made by the script. The speech utterance depicted here is *dat uh was wel goed bevallen toen* ('that uhm was quite well liked then'), totaling 9 syllables, including 'uh'.

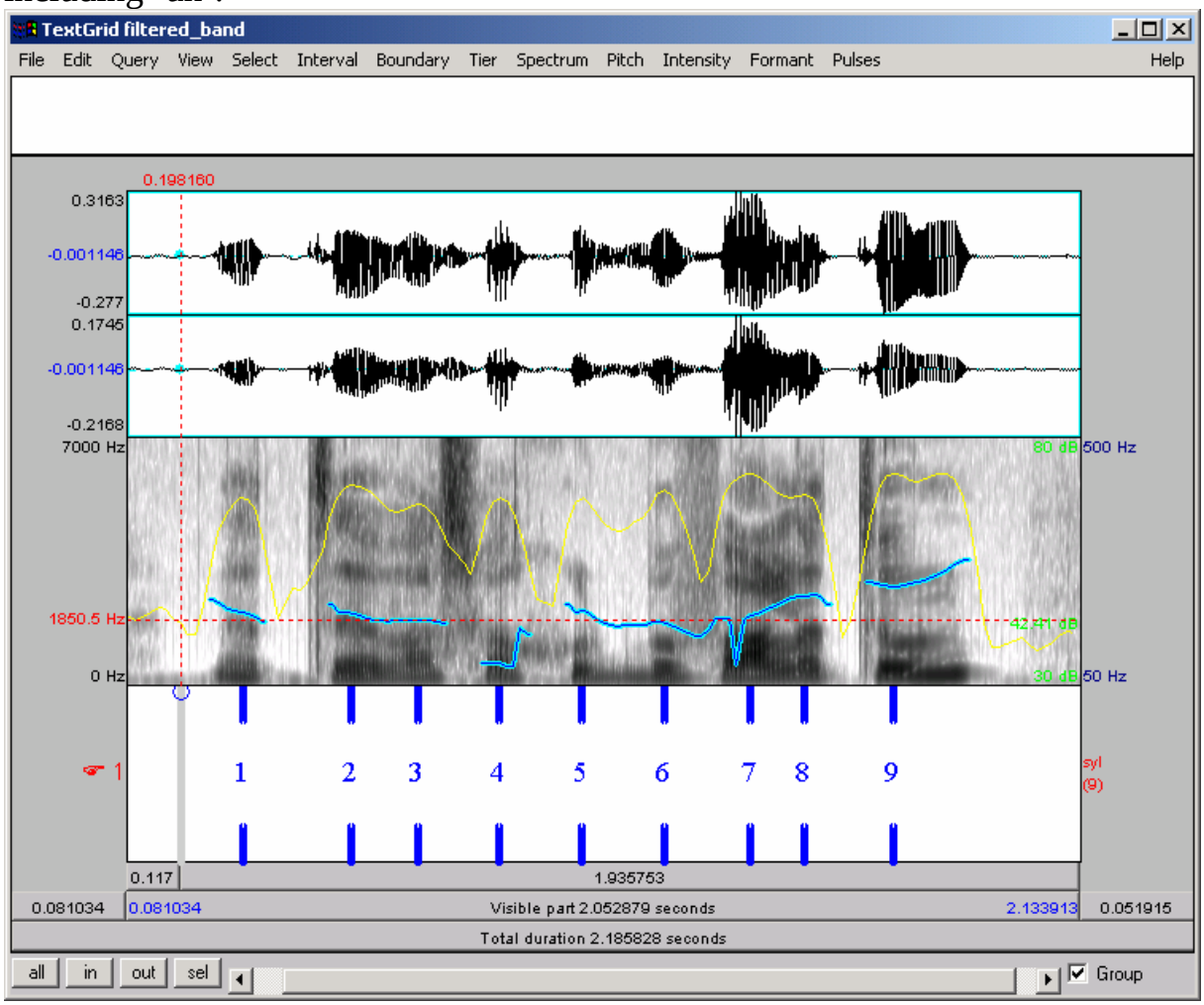



**Figure 1**: part of a speech file in PRAAT with intensity and pitch shown. The points in the tier are the syllable nuclei as detected by the script

## 3 Validation

As a part of the project "What is Speaking Proficiency" (WiSP), conducted at the Amsterdam Center for Language and Communication (ACLC) at the University of Amsterdam, we have collected speech data of 258 participants, 200 non-native speakers of Dutch (with various L1s) and 58 native speakers of Dutch. Each participant performed 8 speaking tasks, resulting in a total of approximately 46 hours of speech. See De Jong, Steinel, Florijn, Schoonen, and Hulstijn (in press) for a description of the speaking tasks and an application of the fluency measures. In order to be able to include measures of fluency in our research, we made two scripts written in PRAAT. The first script automatically detects pauses (a modified version is now incorporated in PRAAT in the *TextGrid (to silences*) button), and the second script automatically detects syllables. The second script is described in this paper. In what follows, we report a validation of the computation of syllables per time unit as generated by the script in two different corpora. First, we randomly selected 50 out of the total of 258 * 8 speaking tasks and measured syllables by hand. This corpus comprised 75 minutes of speech. Secondly, we tested the script on a subset of the IFA-corpus that was comparable to the speaking tasks in the WiSP-study (Van Son, Binnenpoorte, Van den Heuvel, & Pols, 2001). This part of the corpus comprised 125 minutes of speech summed over 8 participants.

### *3.1 Speech data of the Wisp-study*

We counted the syllables of fifty speech files. Pauses longer than 0.4 s were considered possible spurt boundaries. We used all spurts of 5 seconds or more to calculate speech rate, and combined consecutive shorter spurts to get to 5 seconds (excluding pauses). We thus avoided calculating speech rate over very short periods of time. We then automatically detected syllables using the PRAAT-script. Many sound files in this corpus were moderately noisy, therefore we filtered all sounds prior to the syllable measuring, using 100 Hz as the lower edge of the pass band, 5000 Hz as the upper edge of the pass band and 50 Hz as the width of the smoothing region. Measuring peaks in intensity (dB), we used 2 dB above the median intensity per sound file as threshold, and 4 dB as minimum dip between peaks, excluding peaks that are unvoiced.

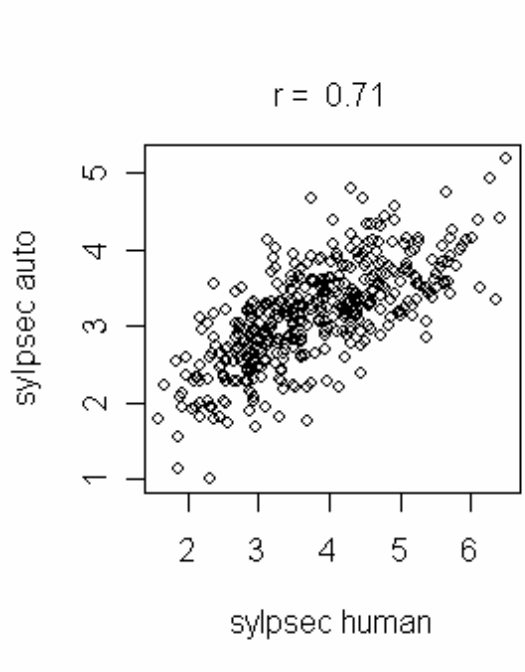



**Figure 2**: Scatterplot of WiSP-speech data of 50 participants, 441 spurts. Number of syllables per second is calculated per spurt by hand and automatically.

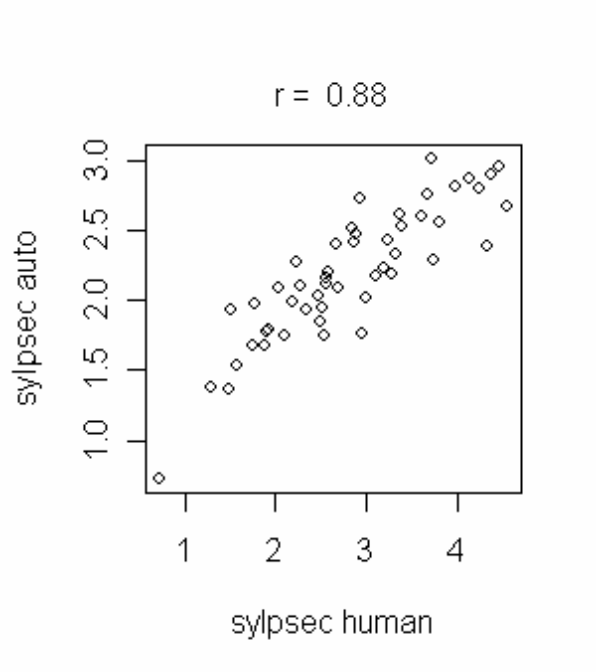



**Figure 3**: Scatterplot of WiSP-speech data of 50 participants. Number of syllables per second is calculated per task (participant) by hand and automatically.

Figure 2 shows the scatterplot of the human and automatic speech rate calculations per spurt; the correlation was .71. For our purposes of comparing speakers and/or tasks, however, we needed a less refined calculation of speech rate. Figure 3 shows the scatterplot when we calculated speech rate over the total speech file: total number of syllables per speech file divided by total speaking time (correlation .88). In other words, the automatically measured speech rate correlates well with human measured speech rate. However, with these data and these parameters, it seems to be the case that the script tends to miss syllables that were actually present. Upon inspection of the TextGrids made by the script, we concluded that the script misses mostly unstressed syllables that were detected by hand.

### *3.2 Speech data of the IFA-corpus*

The IFA-corpus is an open-source database of hand-segmented Dutch speech. Eight participants (4 female, 4 male) performed several speech tasks, ranging from reading aloud lists of syllables to informal story telling. To validate the script on another corpus of Dutch, we selected the three tasks that were similar to the tasks used in the WiSP-study, eliciting (semi-) spontaneous speech. The

three tasks were: informal story telling face-to-face to an "interviewer", retelling of the story previously told, and retelling of a story previously read (Van Son et al., 2001).

For this corpus, we decided not to use a filter, because filtering long sounds takes a lot of time and uses up much computer memory (too much for the computer this script was run on at the time), and because the speech data of this corpus were not as noisy as the above described speech data. As a result, we decided to lower the threshold and minimum preceding dip in intensity. We used the median intensity per sound file as threshold, and 2 dB as minimum preceding dip in intensity. In this corpus, sentences are defined on the basis of pauses as well as syntax, and number of hand-measured syllables could therefore be counted per sentence. As sentences were also defined on syntax, many sentences were very short. Such sentences comprised a single word like "uh", or "en", ("uh" or "and") in mostly beginnings of unfinished sentences.

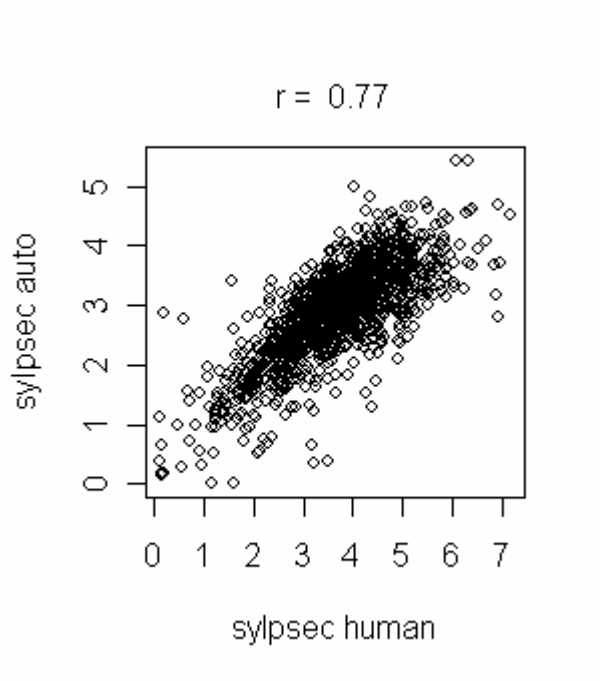



**Figure 4**: Scatterplot of the IFA-corpus, 8 participants, 1171 spurts. Speech rate, number of syllables per second, counted per spurt by hand and automatically.

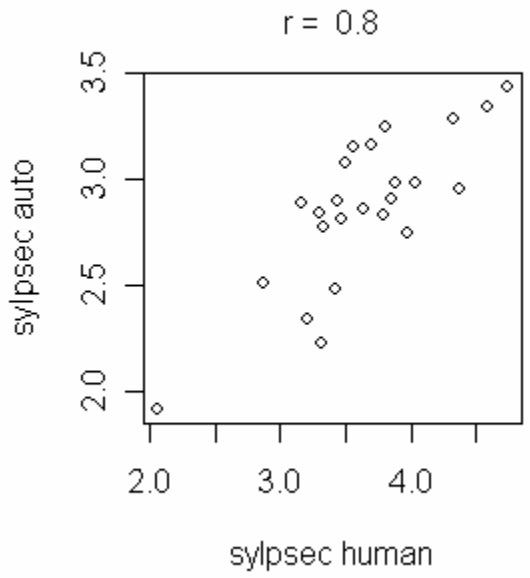



**Figure 5**: Scatterplot of the IFA corpus, 8 participants in 3 tasks. Speech rate, number of syllables per second, calculated per task by hand and automatically. N = 24.

To test the automatic measures against these existing human-made measures, and to be able to compare success of speech rate measures across the two corpora, we redefined spurts in this corpus as stretches of speech (including

pauses) of at least 5 seconds (except when the end of the speech file was reached, in which case the remaining shorter spurt was selected). We then counted the number of syllables using the human transcripts and counted the number of syllables measured automatically for the same time period. In this corpus, we have 8 participants for which human measured information is available in speech tasks quite comparable to those of the WiSP-study. In Figure 4, we show, for all 8 participants, the scatterplot of human measured speech rate per spurt with automatically measured speech rate for that same spurt.

Again, for the purpose of comparing tasks and speakers, we need a calculation of speech rate computed per task. Figure 5 shows the correlation of the 8 speakers in 3 different tasks ($r = 0.8$). As with the speech data of the WiSP study, the script misses syllables that are detected by hand. An inspection of the TextGrids produced by the script, revealed that most of the undetected syllables were unstressed syllables. We think that many of these unstressed syllables might be phonological syllables and therefore detected when measured by hand, but probably not all are also phonetic syllables in the sense that they are present in the signal in any detectable way. Therefore, we may conclude that the algorithm picks up on prominent syllables. As shown by the correlations between human measures and automatic measures, missing such unprominent syllables does not lead to loss of fit. In other words, although the algorithm cannot find all syllable nuclei, it is able to reliably pick up differences between speech rates.

Research by Kormos and Dénes (2004) suggests that in fact it is the number of stressed syllables that correlates best with subjective fluency. Perhaps it is the case that number of prominent syllables better reflects speech rate in the sense that it measures density of content per time unit. Future research is needed to further explore this thesis.

## 4 Conclusion

In this paper, we described a script written in PRAAT that automatically detects syllables in sound files of speech. No transcription of the speech data is necessary to run this script. The script takes sound files as input and writes a TextGrid file with syllable nuclei marked in a point tier. In two validation studies, we found high correlations between human measured speech rate and automatically measured speech rate. Although the script misses (mostly unstressed) syllables that are detected by human judges, the correlations suggest that the algorithm works well in predicting the actual number of syllables. We conclude that for the purpose of measuring speech rate as number of syllables per time unit comparing speakers and tasks, this script suffices.

In second language testing (see, e.g., the speaking rubrics of the TOEFL test as reported on the ETS-website (Educational Testing Service, 2004), second language research (e.g., Kormos & Dénes, 2004), as well as in diagnosing different language and speech disorders (Feyereisen, Pillon, & Partz, 1991; Redmond 2004; Shenker, 2006) fluency is an important factor to take into account. The script described and validated in this paper may be useful to easily and objectively measure speech rate in terms of syllables per second without the need to transcribe speech beforehand.

As yet, it is impossible to directly compare the amount of success of the different syllable measurers available. First of all, other syllable measurers have been developed to detect syllables in spoken English or German, which might be different from detecting syllables in Dutch. Furthermore, the different corpora on which the existing syllable measurers have been tested, have been transcribed by different criteria. Finally, researchers report Pearson correlations for speech rate or for number of syllables *per spurt*. However, comparisons are confounded if spurt length is uncontrolled. For longer spurts, a count of one or two extra or fewer syllables will not result in a large deviation of the calculated speech rate. For short spurts, a count of a single extra or fewer syllable will result in an enormous difference in the calculated speech rate. Future research should take these mathematical issues into account when comparing different methods that automatically measure speech rate. In the present paper, we opted for choosing at least 5 seconds as a constant spurt length. In this way, we were able to compare success in syllable detection across corpora.

## 5 List of references

Boersma, Paul and David Weenink (2007). Praat, http://www.praat.org (Version 4.5.25).

Cucchiarini, Catia., Helmer Strik, and Lou Boves (2002). Quantitative Assessment of Second Language Learners' Fluency: Comparisons between Read and Spontaneous Speech. *The Journal of the Acoustical Society of America, 111*(6), 2862-2873.

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

De Jong, John H. A. L. and Lieneke W. Van Ginkel (1992). Dimensions in oral foreign language proficiency. In Verhoeven, Ludo T. and John H. A. L. De Jong (eds.), *The construct of language proficiency: Applications of psychological models to language assessment.* Amsterdam: John Benjamins, 187-205.

De Jong, Nivja H., Margarita P. Steinel, Arjen F. Florijn, Rob Schoonen, and Jan H. Hulstijn (2007). The effect of task complexity on fluency and functional adequacy of speaking performance. In Van Daele, Siska, Alex Housen, Michel Pierrard, Folkert Kuiken and Ineke Vedder (eds.), *Complexity, Accuracy and Fluency in second language Use, Learning and Teaching.* Brussels: Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, 53 – 63.

De Jong, Nivja H. and Ton Wempe (2007). *Praat script speech rate*. Retrieved December 10th, 2007, from http://home.medewerker.uva.nl/n.h.dejong/bestanden/praatscriptspeechrate.html

Educational Testing Service (2004). *iBT/Next Generation TOEFL Test. Independent Speaking Rubrics*. Retrieved Decmber 10th, 2007, from http://www.ets.org/Media/Tests/TOEFL/pdf/Speaking_Rubrics.pdf

Feyereisen, Pierre, Agnesa Pillon, and Marie-Pierre de Partz (1991). On the Measures of Fluency in the Assessment of Spontaneous Speech Production by Aphasic Subjects. *Aphasiology, 5*(1), 1 - 21.

Hunt, Andrew (1993). Recurrent neural networks for syllabification. *Speech communication, 13*, 323 - 332.

Kormos, Judit and Mariann Dénes (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*(2), 145 - 164.

Mermelstein, Paul (1975). Automatic segmentation of speech into syllabic units. *J. Acoust. Soc. Am., 58*(4), 880 -883.

Pfau, Thilo, Robert Faltlhauser, and Günther Ruske (2000). A Combination of Speaker Normalization and Speech Rate Normalization for Automatic Speech Recognition. *Proc. of ICSLP 2000, Peking, China, 4*(362-365).

Pfau, Thilo and Günther Ruske (1998). Estimating the speaking rate by vowel detection. *Acoustics, Speech, and Signal Processing (ICASSP 2005 Proceedings), 2*, 945-948.

Pfitzinger, Hartmut R. (1999). Local speech rate perception in German speech. *Proc. of the XIVth Int. Congress of Phonetic Sciences, 2*, 893--896.

Redmond, Sean. (2004). Conversational profiles of children with ADHD, SLI and typical development. *Clinical linguistics & phonetics, 18*(2), 107 - 125.

Tavakoli, Parvaneh and Peter Skehan (2005). Strategic planning, task structure, and performance testing. In Ellis, Rod (ed.), *Planning and Task Performance in a Second Language*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 239 - 276.

Van Son, Rob J. J. H., Diana Binnenpoorte, Henk van den Heuvel, and Louis C.W. Pols (2001). The IFA Corpus: a Phonemically Segmented Dutch "Open Source" Speech Database. *EUROSPEECH 2001*, 2051-2054.

Verhasselt, Jan P. and Jean-Pierre Martens (1996). A fast and reliable rate of speech detector. *Spoken Language, (ICSLP 96 Proceedings), 4*, 2258-2261.

Contact information:
Nivja de Jong
nivja.dejong@uva.nl
http://home.medewerker.uva.nl/n.h.dejong

# Investigating the acoustic effect of the descended larynx with articulatory models*

**Bart de Boer**
*ACLC, Universiteit van Amsterdam*

*It has been proposed that the low position of the human larynx (compared to other apes) is necessary for producing distinctive articulations, and that it therefore evolved for the purpose of speech. This idea, however, is controversial. Other animals with low larynges have been described, and speech is also possible without using the full range of possible articulations. The role of the descended larynx has been previously investigated with computer models, but these have produced contradictory results. Here it is proposed that for investigating the role of the descended larynx, articulatory constraints must be taken into account, and therefore computer models must be articulatory models. In this paper a strongly simplified model, as well as three more realistic models are investigated for the effect of larynx height. A short study of human data has also been done. It is found in all cases that a vocal tract with a vertical section that is approximately equally long or slightly shorter than the horizontal section performs best. This corresponds to the anatomy of the female vocal tract. An evolutionary interpretation of these observations could be that the female vocal tract has evolved to be optimal for speech, while the male vocal tract has also evolved under another pressure, most likely size exaggeration.*

## 1 Introduction

Did the human larynx descend because of increased articulatory flexibility? This question has long been debated in the study of the evolution of language (DuBrul, 1958; Fitch, 2000; Lieberman & Crelin, 1971; Negus, 1938). Some researchers suggest that the human descended larynx is an evolutionary adaptation to speech (Lieberman & Crelin, 1971), while others suggest that it is

* The author was funded by the NWO vidi project "modeling the evolution of speech". The author wishes to thank Rob van Son and Louis Pols for critical evaluation of a former version of this manuscript.

the result of other factors, such as bipedal upright walking (Aiello, 1996; DuBrul, 1958) or size exaggeration (Fitch, 2000; Ohala, 1984). Advocates of the adaptive role of the descended larynx in speech point out that it creates room in the pharyngeal region, thus turning the single (oral) cavity vocal tract of other primates into a vocal tract with two controllable cavities: the oral cavity and the pharyngeal cavity. Opponents, apart from proposing other reasons for the descended larynx, put forward that there are many species with descended larynges which do not have enhanced articulatory abilities (Fitch & Reby, 2001), that there are modern human languages that do not use the full range of possible articulations (e. g. Choi, 1991; Ladefoged & Maddieson, 1996, pp. 286-288), and that a descended larynx is not necessary for producing the range of sounds that humans can make (e. g. Boë *et al.*, 2002).

Attempts have been made to build computer models of the acoustic abilities of vocal tracts without a lowered larynx. This is often done together with attempts to reconstruct an ancestral vocal tract, although these are really two logically independent questions. The original example of such work is that by Crelin and Lieberman (Lieberman & Crelin, 1971; Lieberman *et al.*, 1972; Lieberman *et al.*, 1969) who modeled Neanderthal, chimpanzee and rhesus monkey vocal tracts, while more recently Boë and colleagues have made a Neanderthal model (Boë et al., 2002). Similarly, Carré and colleagues (Carré *et al.*, 1995) have made a more theoretical model investigating what vocal tract configuration are needed for producing as distinctive signals as possible. The outcomes of these efforts are contradictory: Crelin and Lieberman find that a lowered larynx is needed and Neanderthals did not have it, and conclude that they were not capable of modern speech. Boë *et al.* find that a lowered larynx is not needed and conclude that Neanderthals were capable of modern speech. Carré *et al.* find that a lowered larynx *is* needed, but they do not take a position on whether Neanderthals had one or not.

Part of the discussion focuses on the reconstruction of the Neanderthal vocal tract. Lieberman's reconstruction has been criticized and more recent reconstructions tend to propose a more human-like shape of the Neanderthal vocal tract (Arensburg *et al.*, 1989; Houghton, 1993; Schepartz, 1993). Still, as no fossilized Neanderthal vocal tract has been found, no definitive conclusions can be drawn. In any case, the question of whether Neanderthals had a vocal tract similar to that of modern *Homo sapiens* is independent from the question what the function of a lowered larynx is.

In the debate about the function of a lowered larynx, there is no disagreement about the basic acoustics: every researcher agrees that a vocal tract with two independently controllable cavities is needed to produce a maximally distinctive set of speech sounds. Rather, the debate focuses on whether a vocal tract that is anatomically like a modern human vocal tract is needed or whether,

with sufficient (voluntarily) articulatory control, an ordinary primate vocal tract can produce the same range of articulations as a modern human vocal tract. Simplifying the debate, one could say that Lieberman et al.'s papers stress the importance of anatomy (most recently made by Lieberman (2006) in reference to Riede et al.'s (2005) model of a Diana monkey), whereas Boë et al.'s papers stress the importance of articulatory control, and propose that sufficient control can overcome the limitations of anatomy. Carré et al.'s contributions do not take a strong position in this debate, but stress the necessity to have two independently controllable cavities, but leave in the middle whether this is achieved through better articulatory control or through specialized anatomy.

In this paper the role of anatomy (in this case, the larynx position) versus articulatory control is investigated with a number of simplified models of the human vocal tract. The question underlying the research is: do differences in anatomy influence the ability of a vocal tract to produce different speech sounds, or are any differences in anatomy offset by articulatory control? First, the effect of larynx height on the range of acoustic signals that can be produced is studied in a highly simplified model of the vocal tract. Second, the articulatory abilities of more realistic models of the human vocal tract are studied. These models are based on the male and female vocal tracts, and it is investigated whether they show the same differences as were found in the simplified model. Finally, it is investigated whether a similar effect can be detected in real human data.

## 2 Basic Methods

As the most important aim of this paper is the theoretical investigation of the role of larynx position on the range of signals that can be produced, and as we cannot independently control larynx height in human subjects, computer models of the vocal tract were used. Another reason to use a computer model is that analytic approaches can only provide limited insight in the case under study. The acoustic effects of certain changes in configuration can be calculated analytically (such as the effect that the resonance frequency of a tube increases if the length is reduced). However, here it is investigated what the effect is on the *range* of signals that can be produced when the way a tract can be *deformed* is changed. This is a complex manipulation of boundary conditions and outside the abilities of ordinary mathematical analysis. Computer modeling is therefore the right tool to study this question.

The experiments presented in this paper are based on the use of geometric articulatory computer models. Such models represent the vocal tract as a number of geometric shapes that can be manipulated with articulatory parameters. The geometric shapes correspond directly to parts of the vocal tract, such as the pharynx, the tongue body, the palate, etc. The articulatory parameters can be

mapped straightforwardly to muscle actions. Such models are the closest approximations one can make to the actions of real vocal tracts. The models used in this paper are based on Mermelstein's (1973) model. His model is a 2-dimensional model of the mid sagittal cross section of the vocal tract, as well as a model of the area of the three-dimensional cross sections at different positions in the vocal tract. Another geometrical model is Goldstein's (1980) model. Although this model has been designed to be able to model male and female vocal tracts, it is not possible to keep the upper vocal tract constant while changing the position of the larynx. Therefore the Mermelstein model was used and modified for the research presented here.

Generating a signal on the basis of an articulatory model involves three steps. The first step is to calculate the 2-dimensional mid sagittal outline of the vocal tract for a given set of articulatory parameters. The second step consists of converting this two-dimensional outline in a function giving the cross-sectional area of the vocal tract at each point along its length. The third step consists of calculating the acoustic properties of a tube with this particular area function. When the vocal tract is modeled as a series of concatenated lossless tubes, this last step can be performed by straightforward application of standard acoustics (Fant, 1960; Flanagan, 1965, section 3.2). The first two steps, however, depend on the anatomy of the vocal tract. Details of the different models will be given in the sections below.

Given an articulatory model, the question rises of how to explore and measure its articulatory abilities. The signal that would be produced by an individual articulation of a given model can be calculated from the lossless tube approximation mentioned above. From this articulation, the position of the resonance peaks can be calculated. Only the first and the second resonances (formants) were used in the analysis. It is well established that although three formants are needed to make all differences between possible vowel signals, the first two formants are the most important cues for establishing vowel quality. Furthermore, the first two formants have been widely used as the basis of perceptual acoustic space. This has not only been the case in research into the evolution of speech (Boë et al., 2002; Carré et al., 1995; Lieberman & Crelin, 1971), but also in research into perception (e. g. Peterson & Barney, 1952) or acquisition of speech (e. g. Kuhl *et al.*, 1997; Kuhl & Meltzoff, 1996).

Every articulation can therefore be considered to result in a point in a two-dimensional acoustic space. The comparison between two different articulatory models then boils down to a comparison of the extent of the areas in acoustic space covered by these two models. Potential measures of this extent are the total area, or measures of its maximal diameter. Here, the measures that are used are the area of acoustic space covered by the articulations a model can make, and the difference between the maximal and minimal formants generated by the

model. The area covered by the articulatory model was calculated by dividing the acoustic space in a grid of squares of $0.5 \times 0.5$ Barks, and counting the number of tiles that had at least one acoustic signal in them. The number of tiles can then be converted to acoustic area by multiplying with the area of a tile ($0.25$ Bark$^2$).

The procedure for exploring the range of possible articulations is inspired by, but not exactly equal to the idea of Maximal Vowel Space as defined by (Boë *et al.*, 1989). It consists of generating a large number of articulations, and calculating what area of acoustic space they cover. It might appear that a systematic exploration of every possible combination of articulatory parameters would be most straightforward. The continuous ranges of articulatory parameter values could be divided into a number (say 10) of equally spaced values, and all possible combinations explored. This approach suffers from two problems, however. The first is that the number of articulations to be explored rises exponentially with the number of articulatory parameters. With 3 articulatory parameters, a thousand articulations must be explored, but with 6 parameters, a million need to be explored. This is infeasible. The second problem is that, due to the discretization of the parameter range, a bias might be introduced. It is quite possible (even likely) that articulations resulting in extreme values of the signal are missed.

A better approach, but at first counterintuitive one, is therefore to generate a large number of random articulations. This is called a Monte Carlo approach in computer science (Metropolos & Ulam, 1949). This approach does not suffer from sampling biases. An added advantage is that the procedure can be repeated a number of times, and the spread of the results be used to get an indication of how well the space is sampled. The "systematic" approach, on the other hand, would always give the same value, and therefore no idea could be obtained of how well the space is sampled.

Only valid articulations were used for calculating the acoustic area. Valid articulations are articulations were the articulators do not intersect. The ranges of the articulatory parameters were selected to be physiologically plausible, but it still remains possible that a combination of articulatory values results in parts of the vocal tract intersecting with each other. This was automatically detected when calculating the value of the cross section areas of the vocal tract. When the diameter was less than zero, intersections occurred, and such articulations were discarded. Only articulations where all cross sections have an area of at least 0.3 cm$^2$ were considered. Given typical airflow rate and a constriction length of 5 cm, smaller areas would cause turbulence (the Reynolds number would be over 2000), and therefore would result in fricatives or fricative vowels. By using this minimal area, the acoustic simulations were therefore limited to airflow without turbulence. Repeating the experiments with 0.1 cm$^2$ and 0.5 cm$^2$ did not

change the qualitative results (although the absolute size of the acoustic space obviously changed).

Finally, the values obtained for the extent of the acoustic space depend on the representation of the formant values. Basic acoustic theory shows that shorter vocal tracts result in higher formants. All else being equal, shorter tracts would therefore result in apparently larger extents in acoustic space. In order to compare the abilities of vocal tracts with different geometries, they would therefore have to be normalized to the same length. It turns out that taking the logarithm of the formant frequencies gives the same result. Furthermore, it turns out that human perception is to good approximation logarithmic as well. Weber's law of perception (Weber, 1834) states that the just noticeable difference between two signals is proportional to the value of these signals. This means that the size of the just noticeable difference is constant when taking the logarithm of signals. In other words, differences between the logarithms of signals give a reasonable measure of their perceptual distance, independent of the actual value of these signals.

A problem occurs at low frequencies, for which Weber's law does not hold completely. A scale that takes this into account is the Bark scale. This scale is also used by other researchers in the field (e.g. Boë et al., 2002). It is also prudent from an evolutionary point of view to use a perceptually accurate scale, as there are indications that perception of speech was already similar for Neanderthals and *Homo sapiens* (Martínez *et al.*, 2004) and it appears as if the basics of perception are much older than any differences in vocal anatomy (Smith & Lewicki, 2006). The exact relation between Hertz and Bark was adopted from (Schroeder *et al.*, 1979; Schwartz *et al.*, 1997) and is as follows:

$$F_{Bark} = 7\sinh^{-1}\left(F_{Hertz}/650\right)$$

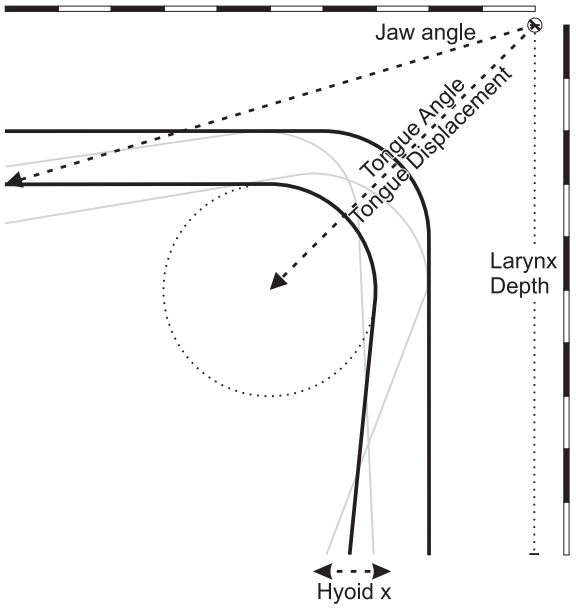
Repeating the experiments and making measurements in either the Mel using the ordinary logarithm of frequency did not result in any qualitative differences (although the exact numerical value and the significances did change).

## 3 The Simplified Model

The simplified model is a stripped-down version of Mermelstein's model. The details of Mermelstein's model are given in appendix A, so only the differences between the simplified and the original model will be discussed. The simplified model does not model the exact anatomy in the region of the lips, nor does it model the exact anatomy in the region of the larynx. This is done to keep the model as simple and symmetric as possible, so that the influence of larynx

height is the determining factor of the model's behavior. Furthermore, the only articulatory motions that were modeled were the motion of the jaw (caused by the mylohyoid and masseter muscles), the motion of the tongue body (both tongue displacement and tongue angle, caused by the styloglossus, genioglossus and hyoglossus muscles) and the horizontal motion of the hyoid. The tract was terminated at the mouth by a vertical plane at a constant position, instead of the more complicated plane that is used in the Mermelstein model. Also, some of the dimensions of the model were simplified somewhat. Finally, cross-sectional diameters are converted to cross sectional areas in the same way everywhere by squaring the value of the diameter.

**Figure 1:** The simplified model. The outline of the model is shown in bold black lines. The articulatory parameters are shown as dashed arrows. The circle that is the basis of the tongue contour and the larynx depth are given as dotted lines. Two potential articulations (different from the rest position) are shown as thin grey lines. For scale, horizontal and vertical bars of 10 cm length are given.

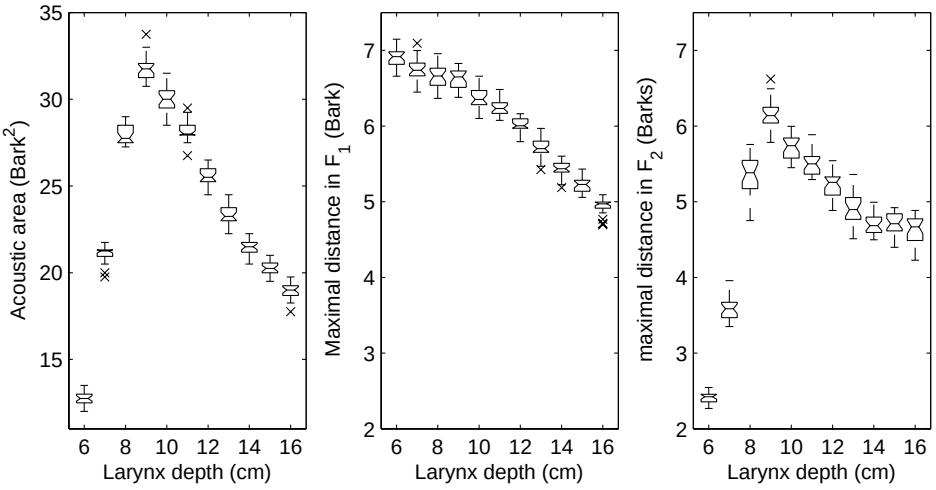



The model is illustrated in figure 1, which can be compared to the more realistic model in figure 3. The ranges of the articulatory parameters are the same as for the Mermelstein model, and are given in table 3 in appendix A. The ranges were determined on the basis of what appeared anatomically plausible, and on the basis of what values did not result in impossible configurations (e.g. ones that intersected themselves). It should be noted that larynx depth is measured with respect to the jaw joint (as are all measurements in the Mermelstein model). Therefore the actual length of the horizontal tube is 8 cm, while the length of the

vertical tube is equal to the larynx depth minus 2 cm (this is done in deference to Mermelstein's original coordinate system, that was relative to the jaw joint).

Models were investigated with larynx depths ranging from 6 cm to 16 cm in increments of 1 cm. For every larynx depth, 100 000 random articulations were generated (where articulatory parameters were uniformly distributed over their range). These were divided into 25 groups of 4000 articulations. For each of these groups the acoustic area and the ranges of $F_1$ and $F_2$ were calculated. The result is shown in figure 2.

**Figure 2:** The relation between acoustic area (left plot) and maximal distance in $F_1$ and $F_2$ (middle and right plots) and larynx depth in the simplified model. The box plot shows the median (horizontal line) as well as the first and third quartiles (top and bottom of the boxes). The total extent of the data set is indicated by the whiskers, while points that can be considered outliers are shown as crosses. Notches in the boxes indicate statistical significance; if the vertical range of the notches of two boxes do not overlap, their difference is significant at the 5% level.



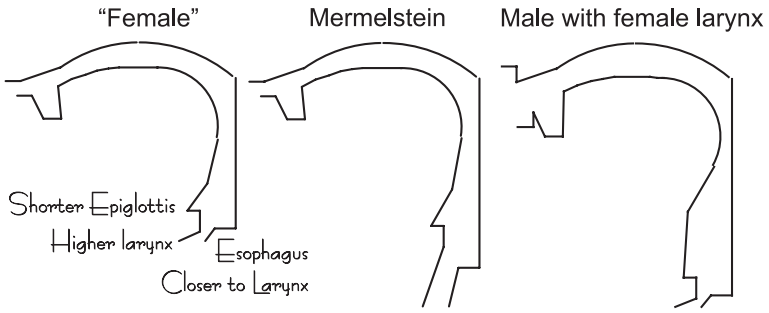

Note that a *smaller* value for larynx depth means a *higher* larynx. It is clear from this figure that the vocal tract that covers the largest acoustic range is the one with a larynx depth of 9 cm. It can also be observed that both higher and lower larynges result in significantly (using the Wilcoxon rank sum test at 5% confidence) smaller reachable areas of acoustic space. There appears to be an optimal larynx depth that is approximately equal to (but in the case of this model slightly smaller than) the horizontal dimension of the vocal tract. It is interesting to note that the difference is mainly due to the inability of models with a lower larynx to produce distinctions in the second formant. As for producing distinctions in the first formants, higher larynges actually appear to be very slightly better than models with medium or low larynges.

## 4 The Realistic Model

Having established that larynx depth influences the signal range of a simplified articulatory model, one can wonder whether this is an artifact of the model, or whether a similar effect obtains in human vocal tracts. Here, no attempt is made to reconstruct a fossil vocal tract. Instead, it is attempted to model the differences in articulatory ability that result from the different larynx positions in human male and female vocal tracts[1].

Mermelstein's model is of the male vocal tract. A reimplementation of his model was therefore used for modeling the male vocal tract. For the female vocal tract, however, his model needed to be modified. As the primary reason for building a female model was to investigate the role of a lowered larynx, as few changes as possible were made to the original model. Using data by Fitch and Giedd (Fitch & Giedd, 1999) as well as Story's data (Story *et al.*, 1996, 1998) it was estimated that the female larynx lays approximately 2.2 cm higher than the male larynx. This corresponds well with the 2.8 cm difference in Goldstein's (Goldstein, 1980) model. It should be noted that when the position of the larynx is mentioned, this is in reference to the larynx at rest. In the Mermelstein model (and in contrast to the simplified model), the larynx *can* move vertically as a result of the motion of the hyoid (caused by the sternohyoid and stylohyoid muscles). This is restricted to ± 0.5 cm.

**Figure 3:** Comparison of the original Mermelstein (1973) model (middle) to the model with the raised, "female" larynx (left) and the mixed model with male position and female shape (right). The differences between the models are indicated in the female model.



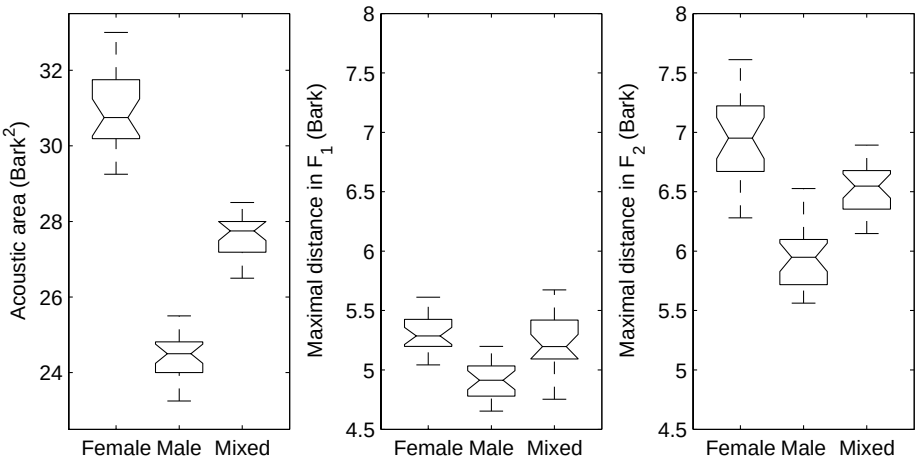

[1] Of course, there is significant individual variation in larynx position and exact vocal tract area function. As the aim of this paper is to investigate the effect of larynx height, only two typical models are investigated, based on data of typical males and females as published in the literature.

There are some other, smaller differences in anatomy as well. Most importantly, the epiglottis is smaller and the esophagus is closer to the larynx in females than in males (Negus, 1949, chapter 11). The larger male epiglottis would not fit in the Mermelstein model with the higher larynx position. In the model used here, the female epiglottis extends upwards 1.7 cm less than the male epiglottis. Finally, also based on Negus's drawings of dissected human larynges, (Negus, 1949, figure 189) the esophagus is modeled to extend 1.3 cm less above the larynx in the female model than in the male model. All these differences are illustrated in figure 3. These differences result in not just a length difference, but also in a different area function, and therefore different volumes of the female and male pharynx. Although this is realistic, it is nevertheless interesting to compare the effect of the lowered larynx alone. A third model was therefore built with a female larynx/epiglottis/esophagus anatomy at the position of the larynx in the male model. This is called the mixed model.

In order to convert the 2-dimensional cross section into a 3-dimensional area function, the same conversion functions are used for the male and the female models. A comparison of area functions derived from MRI-scans of the supralaryngeal vocal tracts of a male subject (Story et al., 1996) and a female subject (Story et al., 1998) articulating the same vowels has shown that most of the difference occurs in the pharyngeal part of the vocal tract. Given that there is considerable inter- and intrasubject variation when producing vowel articulations and that it is unclear whether there are systematic differences between male and female area functions (Soquet *et al.*, 2002) no attempt was made to model the differences in oral vocal tract area function between the male- and female subject. This also minimizes the differences between the models and allows for a purer comparison of the role of the position of the larynx. Details of the model are given in appendix A.

With all three articulatory models, 25 sets of 4000 articulations each were generated. For these articulations, the first and second formant were calculated in Barks as described above. The results are presented in figure 4. All differences are significant with $p < 0.01$, according to the Wilcoxon rank sum test, except for the first formant of the mixed and female models, where there is no significant difference. Apparently the male model, with its lower larynx is able to cover a somewhat smaller range in acoustic space than the female model. This space is even smaller than the space covered by the mixed model, although this model also covers a smaller part of acoustic space than the model with the higher larynx. This is in agreement with the findings of the simplified model.

**Figure 4:** Acoustic area and extent of the first and second formant of the more realistic models. Note that the female model is significantly better in all respects than the male model.




It is important to note that the numbers in figure 4 do not represent a good estimate of the total extent of the acoustic capabilities of the models. Due to the nature of the sampling procedure, the values are always undersampled. However, this has the same bias for every model, and therefore comparisons between models that have been sampled in the same way are valid. In order to get an idea of the total extent of the acoustic space of all models, the values for the complete data set of 100 000 articulations per model can be calculated. These values are given in table 1.

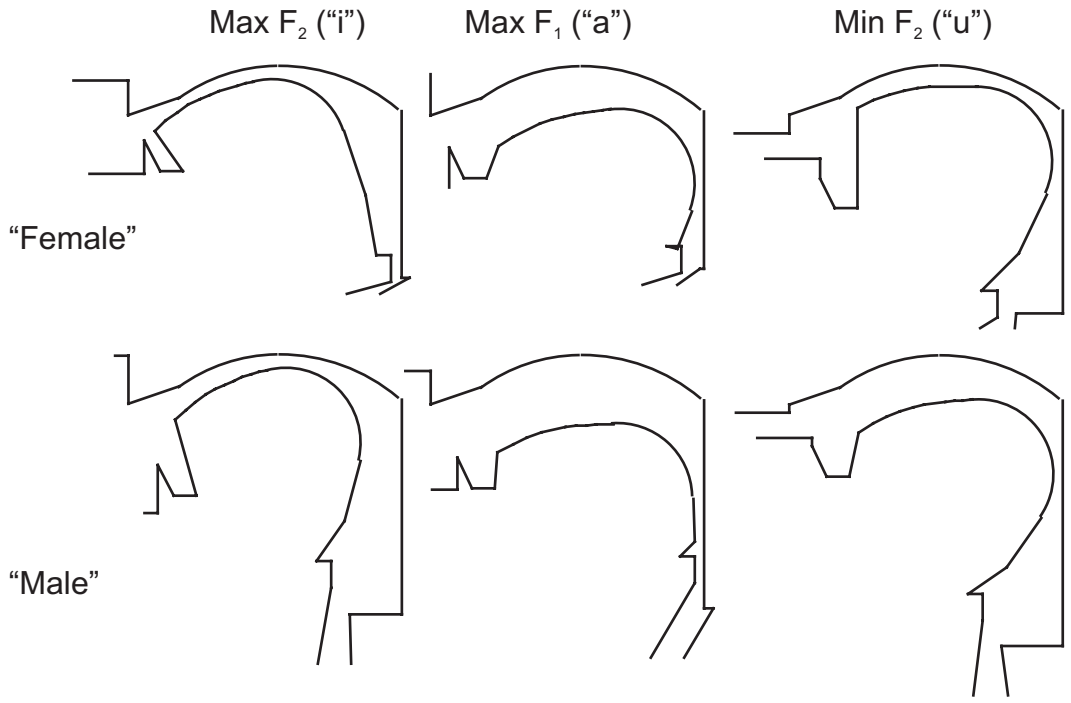
**Table 1** Values of area and maximal extents for the complete data sets.

| | **Female** | **Male** | **Mixed** |
|---|---|---|---|
| Area (Bark$^2$) | 40 | 30 | 36 |
| Max. $F_1$ size (Bark) | 5.8 | 5.4 | 5.9 |
| Max. $F_2$ size (Bark) | 7.8 | 6.9 | 7.5 |

In order to check the validity of the calculated data points, it is necessary to verify whether they correspond to realistic articulations. Of course, this is impossible to do for all 100 000 data points of the data sets. It was therefore only done for the points with the highest and lowest second formant and for the point with the highest first formant for all 100 000 data points. These points correspond roughly to [i], [u] and [a], respectively, and were expected to have the most extreme articulations. They were therefore the articulations that were most likely to be unrealistic. Images of the vocal tract configuration of the male and female models are given in figure 5. The articulations appear to correspond well with the articulations humans make when producing these vowels, with the

possible exception of the female articulation with maximal $F_2$. This articulation appears to have lips that are protruding more than would be the case in a human articulation of [i]. This is most likely an artifact of the undersampling of the available articulatory space by the Monte Carlo method. The articulation with less protruding lips would probably have even higher $F_2$, but it was not sampled. articulations are also plausible[2].

**Figure 5:** Comparison of articulations with maximal $F_2$ and $F_1$ and minimal $F_2$. Note that for both the male and female models, the articulations correspond well with the articulations for [i], [a] and [u]. There appear to be discontinuities in the outline drawn, but this is an artifact of the requirement of the programming environment to use integer values for plotting.




## 5 Real Human Data

All results so far have been obtained with highly stylized models of the vocal tract. It is therefore instructive to check what happens when applying the same measurements to real human data. The effect of lowering of the larynx on the reachable acoustic space can be estimated by comparing the acoustic spread of vowels produced by men with those produced by women.

[2] Hundreds of random articulations were also inspected visually, and none was impossible, although some articulations are at the extreme of what is comfortable, and thus are unlikely to be used in ordinary speech.

Studies of different languages show consistently larger vowel spaces for female articulations than for male articulations (e. g. the data presented in Fant, 1975). Although there is a debate to what extent these results can be explained by behavioral or by anatomical factors (e. g. Diehl, 1996; Goldstein, 1980) the consensus seems to be that anatomy is at least partly the cause.

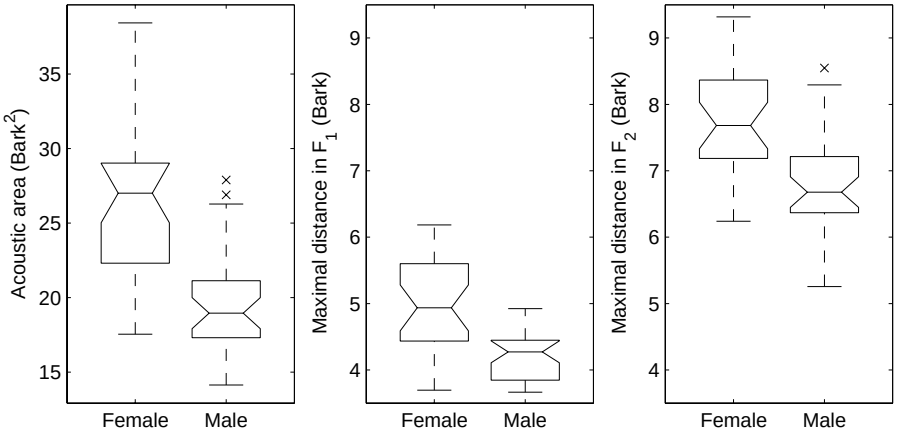
In order to show that the methods used in this paper also work on real data, the classic data set from Peterson and Barney (1952) as reconstructed and made publicly available by Watrous (1991) has been used. This data set is ideally suited for this research, as it contains all data points of all speakers. Comparing acoustic ranges that are assumed to be caused by individual differences in anatomy, cannot reliably be done using averages. The surfaces of the acoustic space (consisting of the first and second formant frequencies in Bark) and the maximal distances in the first and second formant were calculated.

It must be noted that this comparison is different in nature than the one performed on the model-generated data. In the case of the model-generated data, different data sets generated by the same model were compared, whereas in this case, different data sets generated by different speakers are compared. Also, the Peterson and Barney data set only contains 20 vowels per speaker. This makes it impossible to use the same procedure to calculate area as was used in the model study. It was therefore decided to use the *convex hull* to calculate the areas of the vowel spaces of the human speakers. The convex hull is the area that is covered by all linear interpolations between all data points (Cormen *et al.*, 1993, section 35.3). It can also be imagined in the following way: represent every data point by a nail in a flat board, and then stretch a rubber band around the collection of nails. The area inside the rubber band is the convex hull. It was found that in the model-generated data, the area of the convex hull correlates almost perfectly with the area calculated with the grid-based method, although it is systematically higher. This is because the convex hull always includes at least as much or more area than is really covered by a data set. However, when comparing data that is calculated with the same method, this should not be a problem. Data generated by the two different methods should not be compared directly, however.

The different measures of the human data were compared using the Wilcoxon rank sum test. It was found that the female vowel spaces were larger than the male vowel spaces with $p < 0.001$ for all measurements. The measures of male and female vowel spaces are given in figure 6. From this figure it is directly clear that female vowel spaces tend to be somewhat larger than male vowel spaces. Although there might be other reasons for the difference between male and female performance, at least the human data does not contradict the observation from the modeling study that vocal tracts with a larynx in the female position have greater articulatory abilities than vocal tracts with the larynx in the male position. It should be noted, however, that the difference is small, that

there is considerable overlap and that it is not expected that men would have practical difficulties with producing distinctive speech because of this difference.

**Figure 6:** Statistics of the comparison of human female and male data, taken from Peterson and Barney's dataset. Note that absolute values should not be compared with those of the modeling experiments, as the test conditions and the way of measuring acoustic area are different.




## 6 Discussion

In this paper the influence of larynx position on the reachable acoustic space has been investigated. In a strongly simplified computer model of the human vocal tract, it was found that there is an optimal vertical position of the larynx, for which the area in acoustic space covered by the signals that such a model can generate is maximized. In the model the optimal position occurred when the vertical part of the vocal tract was slightly shorter than the horizontal part. The same results were found when more realistic models (based on Mermelstein's model) of the male and female vocal tracts were compared. Here too, the model in which the vertical part of the vocal tract was slightly shorter than the horizontal part could generate a larger repertoire of signals than the model in which the vertical part was longer. A similar difference was found in the Peterson and Barney data set of vowels articulated by male and female speakers.

The modeling results show that articulatory constraints caused by differences in anatomy can influence the range of articulations that can be produced. The simplified models all had the exact same articulatory control, but differed only in the position of the larynx. The more realistic models also all had the same articulatory control, and the mixed model and the female model only differed in the position of the larynx. The male model also had slightly different anatomy in the laryngeal region.

The reason for this difference is that certain configurations of the modeled vocal tract allow for a range of deformations that results in a larger range of acoustic signals, given the way in which the different articulators (most importantly tongue, lips, pharynx and larynx) can be controlled. Apparently, given human-like abilities to control articulation, tracts in which the vertical part of the vocal tract is about equally long as the horizontal part, allow for the greatest range of signals. This agrees with Lieberman et al.(1969)'s analysis.

The obvious question is of course, whether this is also true for the case of human articulation. The results from the comparison of the male and female speakers of the Peterson and Barney data set show a difference that is very similar to the one found between the models of the male and female vocal tracts. Still, it remains possible that this difference is caused by other factors, such as sociolinguistic ones, or biases in measurement. One way to test this is to investigate the way articulatory capabilities change over puberty. The anatomy-is-important hypothesis would predict that boys and girls would be indistinguishable before puberty, but different after puberty. Although research into the effect of development on formant frequencies exists (Lee *et al.*, 1999), their paper only presents averages over age groups, and not individual data points. As the hypothesis presented here is that for males the articulatory range changes over puberty, these averages would be over a multimodal distribution in the critical age group. Therefore the data in (Lee et al., 1999) cannot be used directly. However, given the correct data, the influence of the descent of the larynx in puberty it is empirically testable.

The results do make perfect sense from an evolutionary perspective, however. The female larynx position appears to be very close to the one found to be optimal in the simplified model. This is an argument in favor of the hypothesis (Lieberman et al., 1969) that the human larynx position is optimized for producing as distinctive articulations as possible. The fact that the male larynx is slightly lower than the optimal position could be explained evolutionarily by the fact that this helps to exaggerate size (Fitch, 2000; Ohala, 1984). It has been found that this is important for animals, and it has also recently been found that lower formants help human males to impress other human males (Puts *et al.*, 2006). That the human male larynx is not as low as found in certain animals (Fitch & Reby, 2001) can then be explained by the fact that the male vocal tract needs to remain able to produce a sufficient repertoire of distinctive speech sounds.

Of course this does not mean that human males are necessarily worse at communicating through speech. Indeed, humans can still speak even with deformations of the vocal tract, such as cleft palate, or for that matter, with their mouth full of food. Also, the fact that not all languages use the full extent of possible speech sounds (Choi, 1991; Ladefoged & Maddieson, 1996) indicates

that maximal articulatory capabilities are not essential for modern languages. However, evolution (both cultural and biological) is very good at fine-tuning. Most languages do use the maximally distinctive vowels, for example, and this can be explained as the result of cultural and functional evolution (e. g. de Boer, 2000; Schwartz et al., 1997). As for biological evolution: it is easier to produce distinctive speech when one has the right anatomy. If communication is an important factor in survival, then the larynx position that has the best position for communication will therefore be selected for. The findings of the simplified model indicate that for extremely high larynges (comparable to the chimpanzee vocal tract) small differences in larynx height already make an important difference in useable acoustic space.

The findings of the more complex model and the human data seem to indicate that in human females, the evolutionarily optimal position is near the position that results in the largest range of possible speech sounds, while in human males the evolutionarily optimal position is slightly lower, resulting in lower formants and a more impressive voice.

These results, although certainly not the last word on the role of the descended human larynx, contribute to the debate on whether anatomical constraints are important in understanding the evolution of the vocal tract. They provide an argument that anatomy does matter and that the human vocal tract has a shape that facilitates speech production. It also provides an argument that the female vocal tract is probably the best point of reference when investigating the link between anatomy and distinctive speech.

## 7 Appendix A: Details of the articulatory model.

The articulatory models used in this paper are based on a reimplementation of Mermelstein's (Mermelstein, 1973) model. The reimplementation was based on the description provided in that paper, as well as on the reimplementation by Boersma (Boersma, 1998). However, as the description in Mermelstein's paper was not quite complete, some of the details of the implementation had to be measured from his figures. As the model used here can therefore be somewhat different in its details from Mermelstein's original model, a short description is given. The male and female models are identical, except for a number of parameters describing the shape of the larynx. These can be found in table 2.

The vocal tract outline in the midsagittal plane is approximated by two composite curves consisting of straight lines, circular arcs and a more complex curve to approximate the tongue. The exact shape of these curves is determined by eight articulatory parameters. These are the *x* and *y* position of the hyoid ($x_H$ and $y_H$) the angle of the jaw ($\alpha_J$) the angle and displacement of the tongue ($\alpha_T$

and $d_T$) the angle of the tongue blade ($\alpha_B$) and the protrusion and spread of the lips ($x_L$ and $z_L$). The geometry and the relation of the parameters to the model are illustrated in the left part of figure 7. The minimal and maximal values of the parameters are given in table 3.

**Table 2** The constants used for calculating the outline of the male and female vocal tracts (all lengths in cm, all angles in radians).

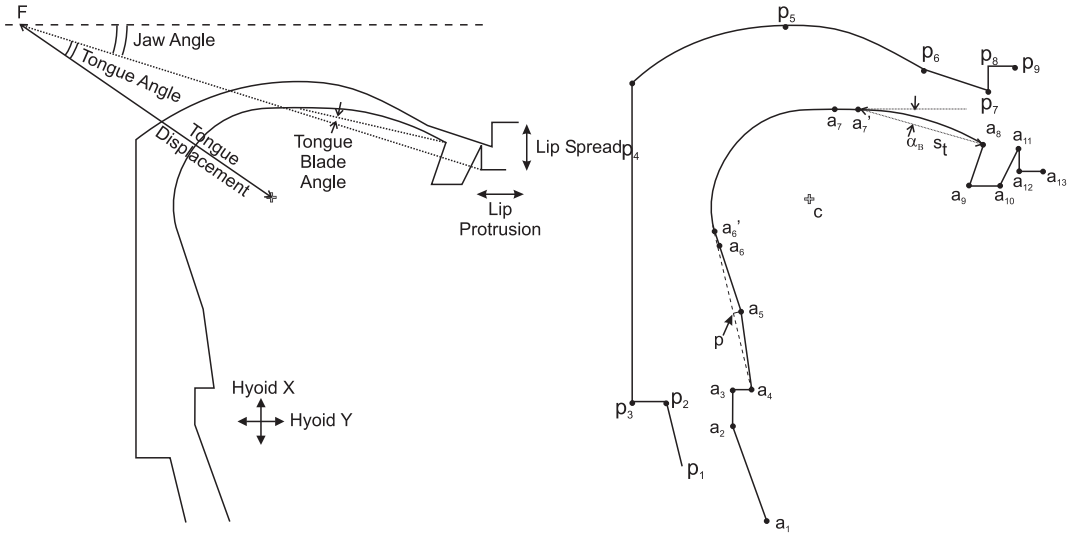
| | **shared** | **male** | **female** | | **shared** |
|---|---|---|---|---|---|
| $H_x$ | 5.6 | | | $H1_x$ | –0.4 |
| $H_y$ | | –7.6 | –7.1 | $H1_y$ | –0.7 |
| $HK_x$ | 0.3 | | | $S$ | 11.3 |
| $HK_y$ | | –2.7 | –1 | $A_i$ | –0.237 |
| $W$ | 0.9 | | | $A_c$ | 0.29 |
| $G_x$ | 4.7 | | | $A_h$ | 1.73 |
| $G_y$ | | –9 | –7.7 | $R_t$ | 2 |
| $R_x$ | 4 | | | $S_t$ | 3.4 |
| $V_y$ | –2.6 | | | $LT2_x$ | –1 |
| $M_x$ | 7.2 | | | $LT_x$ | –0.4 |
| $M_y$ | –1.4 | | | $LT_y$ | –0.8 |
| $N_x$ | 9.87 | | | | |
| $N_y$ | –2.27 | | | | |
| $U_x$ | 11.2 | | | | |
| $U_y$ | –2.7 | | | | |

**Table 3** The minimal and maximal values of the articulatory parameters.

| | min | max |
|---|---|---|
| $x_H$ | –1 | 1 |
| $y_H$ | –1 | 1 |
| $\alpha_J$ | –0.25 | 0.25 |
| $\alpha_T$ | –0.25 | 0.25 |
| $d_T$ | –1.5 | 1.5 |
| $\alpha_B$ | –0.2 | 0.2 |
| $x_L$ | 0 | 1.5 |
| $y_L$ | –1.5 | 1.5 |

The posterior superior outline is defined by nine points $p_1 \ldots p_9$. The anterior inferior outline is defined by thirteen points $a_1 \ldots a_{13}$. The shape of the outlines and the positions of the points are illustrated in the right part of figure 7. The origin of the coordinate system for locating the points is the point around which the jaw rotates, indicated in figure 7 as point *F*.

**Figure 7:** The articulatory parameters (left) and the calculated points on the outlines (right). Note that the posterior/superior outline (with $p_1 \ldots p_9$) and the anterior/inferior outline (with $a_1 \ldots a_{13}$) are only shown in the correct relative positions in the left image.




The points on the posterior/superior outline are calculated as follows:

$$p_1 = \left(x_H/2 + H_x + HK_x - W, y_H + H_y + HK_y\right)$$
$$p_2 = \left(x_H + G_x, y_H + G_y\right)$$
$$p_3 = \left(R_x, y_H + G_y\right)$$
$$p_4 = \left(R_x, V_y\right)$$
$$p_5 = \left(M_x, M_y\right)$$
$$p_6 = \left(N_x, N_y\right)$$
$$p_7 = \left(U_x, U_y\right)$$
$$p_8 = \left(U_x, U_y + z_L\right)$$
$$p_8 = \left(U_x + x_L, U_y + z_L\right)$$

where the constants are given in table 2. All segments are straight lines, except for the segments $p_4$–$p_5$ and $p_5$–$p_6$. These are circular arcs that are horizontal in point $p_5$. The points for the anterior/inferior outline are slightly more complicated to calculate. The first four points are straightforward:

$$a_1 = \left(x_H/2 + H_x + HK_x, y_H + H_y + HK_y\right)$$
$$a_2 = \left(x_H + H_x + H1_x, y_H + H_y + H1_y\right)$$
$$a_3 = \left(x_H + H_x + H1_x, y_H + H_y\right)$$
$$a_4 = \left(x_H + H_x, y_H + H_y\right)$$

points $a_5$ and $a_6$ are calculated with the aid of point $a_6$'. This point is the point where a line from $a_4$ is tangent to the circle describing the tongue body. The tongue body is a circle with radius $R_t$ and center point ($c_x$,$c_y$). The position of the tongue body depends on the motion of the jaw and the tongue as follows:

$$c_x = (S + d_T)\cos(A_j - A_c - \alpha_J - \alpha_T)$$
$$c_y = (S + d_T)\sin(A_j - A_c - \alpha_J - \alpha_T)$$

The point $p$ is midway between $a_4$ and $a_6$'. Point $a_5$ is calculated using a line perpendicular to the line $a_4$–$a_6$'. The length of this line is based on the length of the line $a_4$–$a_6$' as follows: $0.57 \cdot (|a_6' - a_4| - 3.48)$. Note that when the distance is short, $a_5$ moves to the left and when it is long, it moves to the right. Point $a_6$ is now the point on the line tangent to the tongue body, through $a_5$.

Point $a_7$' is the point around which the tongue blade rotates, while point $a_7$ is the point where the tongue blade connects to the tongue body. Point $a_7$' is oriented at a fixed angle with respect to the tongue. As the tongue is rotated by jaw movement, the value of this angle in the absolute coordinate system is changed by jaw movement. The position of point $a_7$' is calculated as follows:

$$a_7' = \left(c_x + R_t \cos(A_j + A_b + \alpha_J), c_y + R_t \sin(A_j + A_b + \alpha_J)\right).$$

The tip of the tongue, point $a_8$ is at a fixed distance from this point, and its position is calculated as follows:

$$a_8 = a_7' + \left(S_t \cos(A_j + \alpha_J + \alpha_T - \alpha_B), S_t \sin(A_j + \alpha_J + \alpha_T - \alpha_B)\right)$$

the starting point of the tongue blade is point $a_7$. This is calculated such that the tongue blade starts tangent to the tongue body, as follows:

$$a_7 = \left(c_x + R_t \cos(A_j + A_b + \alpha_J + \alpha_T - \alpha_B), c_y + R_t \sin(A_j + A_b + \alpha_J + \alpha_T - \alpha_B)\right).$$

It is possible that in this respect the model described here differs slightly from Mermelstein's model, as the starting point of the tongue blade is somewhat unclear in his description.

In order to calculate the final points on the anterior-inferior outline, the motion of the jaw needs to be taken into account. In Mermelstein's model, it is assumed that there is a fixed distance between the top of the lower teeth and the point around which the jaw rotates. The top of the lower teeth is point $a_{11}$ in our model. Its position is calculated as follows:

$$a_{11} = \left(S_j \cos(A_j - \alpha_J), S_j \sin(A_j - \alpha_J)\right)$$

Points $a_9$ and $a_{10}$ are relative to this point:

$$a_9 = a_{11} + \left(LT2_x, LT_y\right)$$
$$a_{10} = a_{11} + \left(LT_x, LT_y\right).$$

Points $a_{12}$ and $a_{13}$ depend on both the position of the jaw and on the spread and protrusion of the lips, as follows:

$$a_{12} = a_{11} + \left(0, -z_L\right)$$
$$a_{13} = a_{11} + \left(x_L, z_L\right).$$

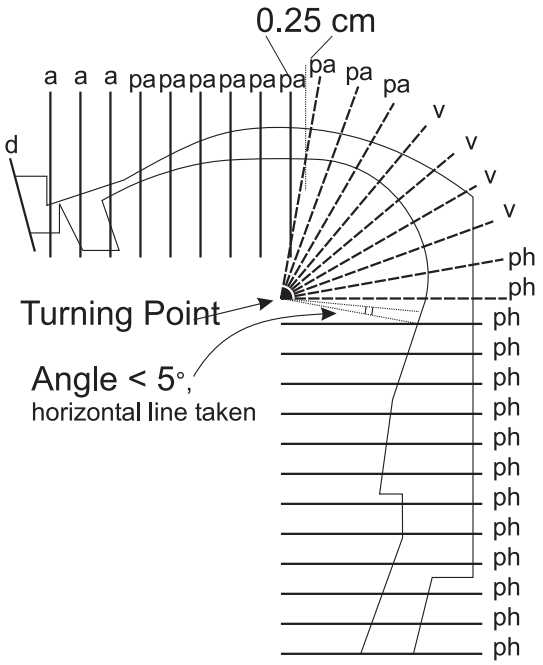
All points are connected by straight lines, except $a_6$ and $a_7$, which are connected by a circular arc with radius $R_t$, and $a_7$ and $a_8$, which are connected by a curve that is quadratic in polar coordinates. The starting radius is the tongue radius, and the ending radius is the distance between the tongue center and point $a_8$. The starting angle is the angle of the line from the tongue center to $a_7$, and the ending angle is the angle between the tongue center and $a_8$. The angle changes linearly, while the radius increases quadratically, according to the following equation:

$$r_t = \left(1-t\right)^2 r_{start} + t^2 r_{end}.$$

These equations and constants give the complete 2-dimensional midsagittal section of the model. On the basis of this section, the area function of the vocal tract can be estimated. In order to do this, a number of cross sections is calculated. In accordance with Mermelstein's model, the cross sections in the laryngeal/pharyngeal section of the vocal tract are horizontal, the cross sections in the uvular/velar area are radial, and in the front part of the vocal tract, they are vertical. This is illustrated in figure 8. The horizontal sections start at the lowest possible point of the larynx, and continue until they reach the vertical coordinate –4.3 cm (relative to the jaw turning point). They are spaced 0.5 cm apart. The radial sections all pass through the *turning point* (7.2, –4.3). The first radial section starts at an angle that is 5º counterclockwise of the line that passes through the turning point and through the intersection between the anterior/inferior outline and the last horizontal section, or, if this would result in an intersection pointing downwards, is taken to be horizontal (still passing through the turning point). Subsequent intersections are taken at intervals of 10º, until they pass the vertical. The vertical sections start at a point that is 0.25cm to the left of the last anterior/inferior intersection of a radial line. They are spaced 0.5 cm apart, and continue until an intersection with either the posterior/superior outline or the anterior/inferior outline is no longer possible.

The last section is taken between point $p_8$ and $a_{13}$. These represent the horizontal extremities of the lips.

**Figure 8:** Sections as used for calculating the cross sectional area in the model. Note that precise positions of the sections depend on the position of the larynx.




As these sections do not accurately follow the orientation of vocal tract, it is necessary to scale them and interpolate them in order to obtain accurate uniform tubes. For this, first the centerline of the vocal tract is calculated. Then for each section, the angle with the centerline is calculated. For all sections except the first and the last this is done by calculating the two angles between the line perpendicular to the section and the two lines connecting the center of the section with the centers of the preceding and following sections. The angle with the centerline is then taken as the average of these two angles. For the first and the last sections, only the angles with the lines connecting to the following or the preceding sections are used. The length of the section is then multiplied with the sine of this angle in order to obtain the cross-sectional length perpendicular to the center line.

With these scaled cross sectional lengths, the areas of the vocal tract at the sections are estimated using Mermelstein's equations. Mermelstein uses different functions converting cross section into area for different regions of the vocal tract. As the model described here is slightly different in its calculation of cross-sections and as it must work for vocal tracts of different lengths, it is possible that the exact extent of the different regions is slightly different for our model than for Mermelstein's. It does follow his model as closely as possible, however. Unfortunately, it is not exactly defined in Mermelstein's paper where

the different regions start and end. In the reimplementation the following criteria were used: for the pharyngeal region, the y-coordinate of the posterior-superior intersection must be less than –3.3, for the velar region the y-coordinate must be less than –1.8. For the palatal region, the x-coordinate of the posterior-superior intersection must be less than 9.7 and for the alveolar region the x-coordinate must be less than 11.2. The rest of the sections are considered to be part of the labial region.

For the pharyngeal region the area is that of an ellipse with one axis increasing from 1.5 cm at the larynx to 3 cm at the upper end. The other axis has the length of the cross section. For the velar region, the area is $2c^{1.5}$ where $c$ is the length of the cross section. For the palatal region, it is $1.6c^{1.5}$. For the alveolar region, the following scheme is used:

$$\begin{array}{rl} 1.5c & c < 0.5 \\ 0.75 + 3(c - 0.5) & 0.5 \le c < 2 \\ 5.25 + 5(c - 2) & 2 \le c \end{array}.$$

Finally, the labial region, is again assumed to be elliptical, with one axis equal to the cross sectional length and the other equal to: $2 + 1.5(z_L - x_L)$.
Areas must minimally be 0.3 $cm^2$ (or the other values specified in the paper), while maximal areas for the different regions are given in table 4.

**Table 4:** Maximal areas in the different regions of the vocal tract (in $cm^2$).

| **Region** | **Max. area** |
|---|---|
| Pharyngeal | 6 |
| Velar | 5 |
| Palatal | 7 |
| Alveolar | 8 |
| Labial | 15 |

The center points of the cross sections are not equally far apart. In order to approximate the vocal tract with uniformly spaced tubes, a linear interpolation of the calculated areas is done at intervals of 0.5 cm. At the end of the vocal tract, two more tubes of 12 $cm^2$ and 30 $cm^2$ are added in order to model the way sound is radiated at the lips. Acoustic properties are calculated as if all sound waves are reflected at the glottis and all waves are radiated at the final tube.

## 8 List of references

Aiello, Leslie C. (1996). Terrestriality, bipedalism and the origin of language. In W. G. Runciman, J. Maynard Smith & R. I. M. Dunbar (Eds.), *Evolution of social behaviour patterns in primates and man* (pp. 269–289). Oxford: Oxford Academy Press.

Arensburg, B., Tillier, A. M., Vandermeersch, B., Duday, H., Schepartz, L. A., & Rak, Y. (1989). A middle palaeolithic human hyoid bone. *Nature, 338*(6218), 758–760.

Boë, Louis-Jean, Heim, Jean-Louis, Honda, Kiyoshi, & Maeda, Shinji. (2002). The potential neandertal vowel space was as large as that of modern humans. *Journal of Phonetics, 30*(3), 465–484.

Boë, Louis-Jean, Perrier, Pascal, Guerin, Bernard, & Schwartz, Jean-Luc. (1989). *Maximal vowel space.* In Eurospeech-1989, 2281–2284

Boersma, Paul. (1998). *Functional phonology*. The Hague: Holland Academic Graphics.

Carré, René, Lindblom, Björn, & MacNeilage, Peter F. (1995). Rôle de l'acoustique dans l'évolution du conduit vocal humain. *Comptes Rendus de l'Académie des Sciences, Série II, 320*(série IIb), 471-476.

Choi, John D. (1991). Kabardian vowels revisited. *Journal of the International Phonetic Association, 21*, 4–12.

Cormen, Thomas H., Leiserson, Charles E., & Rivest, Ronald L. (1993). *Introduction to algorithms*. Cambridge (MA): The MIT Press.

de Boer, Bart. (2000). Emergence of vowel systems through self-organisation. *AI Communications, 13*, 27–39.

Diehl, Randy L. (1996). On explaining certain male-female differences in the phonetic realization of vowel categories. *Journal of Phonetics, 24*(2), 187-208.

DuBrul, E LLoyd. (1958). *Evolution of the speech apparatus*. Springfield (IL): Charles C. Thomas.

Fant, Gunnar. (1960). *Acoustic theory of speech production*. 's Gravenhage: Mouton.

Fant, Gunnar. (1975). Non-uniform vowel normalization. *Speech Transmission Laboratory Quarterly Progress and Status Report, 16*(2–3), 1–19.

Fitch, W. Tecumseh. (2000). The evolution of speech: A comparative review. *Trends in cognitive sciences, 4*(7), 258–267.

Fitch, W. Tecumseh, & Giedd, Jay. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America, 106*(3, Pt. 1), 1511–1522.

Fitch, W. Tecumseh, & Reby, David. (2001). The descended larynx is not uniquely human. *Proceedings of the Royal Society of London Series B - Biological Sciences, 268*, 1669–1675.

Flanagan, James L. (1965). *Speech analysis, synthesis and perception*. Berlin: Springer.

Goldstein, Ursula Gisela. (1980). *An articulatory model for the vocal tracts of growing children.* Unpublished PhD, Massachusetts Institute of Technology, Cambridge (MA).

Houghton, Philip. (1993). Neandertal supralaryngeal vocal tract. *American Journal of Physical Anthropology, 90*, 139-146.

Kuhl, Patricia K., Andruski, Jean E., Chistovich, Inna A., Chistovich, Ludmilla A., Kozhevikova, E. V., Rysinka, V. L., Stolyarova, E. I., Sundberg, U., & Lacerda, F. (1997). Cross-language analysis of phonetic units in language addressed to infants. *Science, 277*, 684–686.

Kuhl, Patricia K., & Meltzoff, Andrew N. (1996). Infant vocalization in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America, 100*(4), 2425–2438.

Ladefoged, Peter, & Maddieson, Ian. (1996). *The sounds of the world's languages*. Oxford: Blackwell.

Lee, Sungbok, Potamianos, Alexandros, & Narayanan, Shrikanth. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of the Acoustical Society of America, 105*(3), 1455–1468.

Lieberman, Philip H. (2006). Limits on tongue deformation - diana monkey formants and the impossible vocal tract shapes proposed by Riede *et al.* (2005). *Journal of Human Evolution, 50*(2), 219–221.

Lieberman, Philip H., & Crelin, Edmund S. (1971). On the speech of neanderthal man. *Linguistic Inquiry, 2*, 203–222.

Lieberman, Philip H., Crelin, Edmund S., & Klatt, Dennis H. (1972). Phonetic ability and related anatomy of the newborn and adult human, neanderthal man, and the chimpanzee. *American Anthropologist, 74*, 287–307.

Lieberman, Philip H., Klatt, Dennis H., & Wilson, William H. (1969). Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates. *Science, 164*, 1185–1187.

Martínez, I., Rosa, M., Arsuaga, J.-L., Jarabo, P., Quam, R., Lorenzo, C., Gracia, A., Carretero, J.-M., Bermúdez de Castro, J.-M., & Carbonell, E. (2004). Auditory capacities in middle pleistocene humans from the sierra de atapuerca in spain. *Proceedings of the National Academy of Sciences, 101*(27), 9976–9981.

Mermelstein, P. (1973). Articulatory model for the study of speech production. *Journal of the Acoustical Society of America, 53*(4), 1070–1082.

Metropolos, Nicholas, & Ulam, Stanislaw. (1949). The monte carlo method. *Journal of the American Statistical Association, 44*(247), 335–341.

Negus, V E. (1938). Evolution of the speech organs of man. *Archives of Otolaryngology, 28*, 313–328.

Negus, V E. (1949). *The comparative anatomy and physiology of the larynx*. London: William Heinemann Medical Books Ltd.

Ohala, John J. (1984). An ethological perspective on common cross-language utilization of f0 of voice. *Phonetica, 41*(1), 1–16.

Peterson, Gordon E., & Barney, Harold L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America, 24*(2), 175–184.

Puts, David Andrew, Gaulin, Steven J. C., & Verdolini, Katherine. (2006). Dominance and the evolution of sexual dimorphism in human voice pitch. *Evolution and Human Behavior, 27*(4), 283–296.

Riede, T., Bronson, E., Hatzikirou, H., & Zuberbühler, K. (2005). Vocal production in a non-human primate: Morphological data and a model. *Journal of Human Evolution, 48*(1), 85-96.

Schepartz, L. A. (1993). Language and modern human origins. *Yearbook of Physical Anthropology, 36*, 91–126.

Schroeder, Manfred R., Atal, B. S., & Hall, J. L. (1979). Objective measure of certain speech signal degradations based on masking properties of human auditory perception. In B. Lindblom & S. Öhman (Eds.), *Frontiers of speech communication research* (pp. 217–229). London: Academic Press.

Schwartz, Jean-Luc, Boë, Louis-Jean, Vallée, Nathalie, & Abry, Christian. (1997). The dispersion-focalization theory of vowel systems. *Journal of Phonetics, 25*, 255–286.

Smith, Evan C., & Lewicki, Michael S. (2006). Efficient auditory coding. *Nature, 439*(7079), 978–982.

Soquet, Alain, Lecuit, V., Metens, Thierry, & Demolin, Didier. (2002). Mid-sagittal cut to area function transformations: Direct measurements of mid-sagittal distance and area with mri. *Speech Communication, 36*, 169–180.

Story, Brad H., Titze, Ingo R., & Hoffman, Eric A. (1996). Vocal tract area functions from magnetic resonance imaging. *Journal of the Acoustical Society of America, 100*(1), 537-554.

Story, Brad H., Titze, Ingo R., & Hoffman, Eric A. (1998). Vocal tract area functions for an adult female speaker based on volumetric imaging. *Journal of the Acoustical Society of America, 104*(1), 471–487.

Watrous, Raymond L. (1991). Currrent status of Peterson-Barney vowel formant data. *Journal of the Acoustical Society of America, 89*(5), 2459–2460.

Weber, Ernst Heinrich. (1834). *De pulsu, resorptione, auditu et tactu: Annotationes anatomicae et physiologicae*. Leipzig: Koehler.

Contact information:
Bart de Boer
ACLC
Universiteit van Amsterdam
Spuistraat 210
1012 VE Amsterdam
tel. +31 20 525 2182
b.g.deboer@uva.nl

# Disapprobation expressions are vocative epithets*

**Boban Arsenijević**
*University of Amsterdam*

*The paper discusses a particular type of disapprobation expressions. These expressions are discussed in Potts & Roeper (2006), where they are referred to as Expressive Small Clauses, and argued to involve the syntactic structure of a bare small clause, and to derive the semantic type of expressions, with the value in the domain of expressivity, instead of the traditional sets of entities and truth-values. The paper argues that these expressions are better analyzed as vocative epithets. Hence, they involve a structure rich enough to include inherent case, they are of the type <e>, and all the pragmatic effects come from the pragmatic integration of this semantically in no way special expressions in the discourse.*

## 1 Introduction

The paper is an extended comment on Potts & Roeper (2006, henceforth: P&R), who propose an analysis for two-member expressive constructions that they call Expressive Small Clauses (ESC), illustrated in (1). P&R divide these constructions to those expressing disapprobation (1a-b) and those expressing incredulity (1c-d).

(1) a. Oh, you fool!
b. You idiot!
c. A: You are an idiot! B: Me an idiot?!
d. A: You are so wrong. B: Me wrong?

I show that their Small Clause analysis of the disapprobation expressions (1a-b) is not correct, and argue that there are two different constructions of this type, one involving vocative epithets, and the other regular predications with extensive ellipsis. I discuss general conclusions that they draw based on their analysis, about language acquisition and the status of expressive meanings.

* The present work is a result of my engagement at the NWO-funded project "The origins of truth and the origins of the sentence".

## 2 Disaprobation ESCs

P&R present ESCs as very simple predications (real bare small clauses), the meanings of which do not correspond to truth values, but instead to some expressive content. They argue that the disapprobation ESCs, as in (1a-b), are closely related to the two word utterances used by children, which are also standardly considered small clauses (Lebeaux 1988, Radford 1990). Their reasoning goes as follows. At the two-word utterance stage, children have not yet acquired higher functional projections, related to the discourse and propositionality, and their linguistic performance does not yet involve a (full-fletched) discourse representation. Hence, children at this stage can only produce bare predications, in the structurally simple form of small clauses. A large number of these small clauses, or – as P&R speculate – maybe even all, are ESCs. In other words, at this stage, children mostly, or only, make utterances with expressive contents: instead of updating a discourse, their expressions simply externalize their thoughts, lexicalizing predications which they want to express. Higher structural projections, specialized for realizing different kinds of meanings related to the discourse, are acquired at later stages. These higher projections, when acquired, become a necessary part of every sentence, and the discourse-related content (assertion, question, topic) that they add blocks pure expressive meanings from realization. The only way for adults to realize expressive meanings is still through ESCs, which therefore continue to live in the language of adults.

P&R's analysis, and the conclusions they draw based on it, has the requirements as formulated in (2).

(2) a) the observed disapprobation expressions do not need discourse for their interpretation; if they did, they would need to project the structure related to the discourse and to have some status in the discourse as well (e.g. of propositions).
b) none of the elements of the observed disapprobation expressions involves case – because structural cases are related to tense (Pesetsky & Torrego 2000) and thus also to the higher functional projections, and inherent cases are untypical for and difficult to motivate in small clauses.
c) expressions with the CP layer cannot bear (pure) expressive meanings; if they could, the preservation of the ESCs in the later stages would be hard to motivate.

Two arguments against their analysis come from their own data. Observe the incredulity expressions in (1c-d), argued by P&R to have expressive meanings. Both examples, as well as many others, involve some focus intonation, and focalization is only interpretable in relation to the discourse. In fact, a large number of foci are directly related to different expressive meanings (e.g. being shocked). The other two conditions will be the topic of the next section. Moreover, observe a variant of what would be a typical ESC in P&R's terms, in (3). This expression involves the clitic *ž(e)*, which is argued to be related to the discourse and generated in a functional projection – something that is supposed to be absent from P&R's structure of an ESC.[1]

(3) Ah, ty ž voryuga! Russian
Ah you.NOM PTCL stealer.NOM
'Ah, you stealer!' (Marina Dyakonova, p.c.)

The second argument relates to the discourse-dependence of the expressive power of the predicate used in the expression. P&R note that even nouns like *waiter* can be used with the expressive power of those like *fool* or *idiot*. In fact, virtually any noun can be used for this purpose under the proper pragmatic circumstances.

(4) a. You are trying to make allies again. You politician!
b You broke a glass again. You elephant!
c. You just step on things and break them. You tank!

As the examples in
(4) show, all we need to assign a noun expressive potential is a proper context, which may be provided in the discourse. But then disapprobation expressions of this type are sensitive to the discourse, at least with respect to the expressive power of the predicates they involve, contra the requirement in (2a). This points to one deeper aspect of the problem: if these expressions were independent of the discourse, we would in fact expect them to be solipsistic to some extent. This means that a large number of such expressions should be uninterpretable by the addressee (like children's two word utterances often are), because the addressee's ability to interpret depends on the discourse salience of the expression. But this is not he case, indicating that these expressions do have an access to the discourse.

[1] I am grateful to Marina Dyakonova for pointing out to me this construction.

## 3 Vocative epithets

P&R provide data illustrating disapprobation expressions from several languages. Russian examples, for instance, differ from the English ones in not realizing overtly what P&R call the subject of the small clause (*you* in all the examples in (1)), while Afrikaans may realize a contracted copula between the two elements of the ESC, resulting in a construction that looks like a full predication rather than a small clause. Nevertheless, the authors consider there to be only one type of disapprobation ESCs. In this section, I provide evidence from Serbo-Croatian that there are at least two different disapprobation constructions among the examples P&R discuss. I argue that the analysis by P&R is not correct and offer a theoretically and empirically better one.

(5) a. Andries, jy's 'n idioot! Afrikaans
Andries you're an idiot
b. Idiotka! Spoken Russian
idiot.FEM

Based on their Afrikaans data, in which a nominal in vocative appears with the ESC, Potts and Roeper observe a possible alternative analysis, namely that what they call disapprobation ESCs are, in fact, vocatives. However, they discard this line of explanation, pointing that vocatives also appear in non-disapprobation contexts, as in (6), and that, in English, vocative is not used for self-reference.

(6) You, waiter! Could you bring me a glass of water, please?

Note that the first argument is based on the completely arbitrary assumption that vocatives can be used in only one type of constructions. However, observe that, were *you* in (1a-b) a vocative, there would be two different embeddings of vocatives in the two types of use. When used without the expressive function, the two nominals in vocative, *you* and *waiter* in (6), are prosodically separated. When used as disapprobation expressions (like in (1a-b), or even (6) if *waiter* is understood with a sense of disapprobation), the two elements make one prosodic word, with the stress on the predicate: *you-IDIOT*, *you-FOOL*, *you-WAITER*. This suggests that *you* in these examples is a second person vocative clitic (almost a vocative affix of the predicate), and not a full pronoun.

One may challenge the presented analysis by asking why then other vocatives in English, like *waiter* in (6), do not need the vocative clitic. The answer that I would suggest is that regular vocatives phonologically realize the second person feature by means of prosody: shouting them with a particular intonation marks their vocative nature, including the second person. The disapprobation expressions under discussion are, however, not in the default vocative use of attracting the addressee's attention. The predicative content

introduced by the epithet requires a special kind of intonation, which blocks the default vocative prosody. Therefore, the clitic needs to be inserted in order to realize the second person feature of the vocative, which, unlike the attention-attracting feature, needs to be preserved in the expressive use.

The second argument was that the English *you* in vocative is not used for self-reference. However, native speakers judge fine examples like in (7).

(7) a. And $I_i$ said to $myself_i$: $honey_i$, $you_i$ should definitely learn how to make pasta.
b. $I_i$ realized $I_i$ was snoring, and $I_i$ thought: hey $you_i$! wake up! what are $you_i$ doing?!

Finally, even if the pronoun *you* in vocative is not used for self-reference, the clitic *you* in the ESCs may well be, taken that the suggested analysis is correct.

Support for the vocative epithet analysis also comes from the way these expressions are treated in communication. Observe the possible dialog in (8), where the ESC is referred to as calling someone a name, the default function of the vocative case.

(8) A: What are you doing?! You idiot!
B: Did you just call me an idiot?

The crucial piece of evidence in favor of the vocative analysis is empirical, and comes from Serbo-Croatian, where vocatives are used for both different functions. The exact Serbo-Croatian counterparts of the English expressions in (1a-b) are given in
(9).

(9) a. E, budal-o! Serbo-Croatian
oh fool.VOC
b. Idiot-e!
idiot.VOC

As obvious, in both cases, the noun is in the morphological vocative case. Moreover, just like in Afrikaans, this kind of expression is quite often used with a name in vocative, but then with a prosodic pause between the two vocatives. One of the vocatives is here used with the default vocative function and the other as a disapprobation expression.

(10) a. E, Boban-e, idiot-e! Serbo-Croatian
oh, Boban.VOC idiot.VOC

b. Boban-e, budal-o!
Boban.VOC fool.VOC

A straightforward analysis is that the expressive part (*idiote*, *budalo*) is an epithet in vocative. Using an epithet to call someone realizes the implication that we consider the addressee a bearer of the property denoted by the epithet. When it appears with the name in vocative, it is probably an appositive, as indicated by the prosodic pause. We may formulate this conclusion as in (11).

(11) ESCs are properly analyzed as epithets taking the vocative case.

In the theoretical view, an important advantage of this analysis is a unification of explanations for the behavior that disapprobation expressions and vocatives show with respect to embedding. As P&R show, the English type of disapprobation expressions has the peculiar property of not embedding into other syntactic structures, or allowing more complex structures than an NP to embed inside them. The same behavior is characteristic of vocatives: they present nominal expressions which do not participate in other syntactic structures. The present paper does not have the ambition of explaining vocatives, but whatever explanation someone proposes for their behavior, it should hold for disapprobation expressions as well.

In addition to offering a better analysis for the English and Serbo-Croatian disapprobation expressions, this section also demonstrates that the Serbo-Croatian ones fail the requirement in (2b): they do involve inherent case, namely vocative, and hence are not real small clauses.

## 4 One or more disapprobation expressions

So far, I have shown, contra P&R, that there are disapprobation expressions that are properly analyzed as vocative epithets (the ones in Serbo-Croatian), and suggested that the same vocative construction is used in other languages as well, including English. It is, however, still possible that there are two or more different constructions used for the same purpose in different languages. In this section, I demonstrate that this is indeed the case. I preserve, nevertheless, the stand that the English construction is just like the one in Serbo-Croatian, a vocative epithet, with the analysis suggested in the preceding section (*you* being a vocative marker, not a real pronoun).

P&R consider all disapprobation expressions in different languages to have the same underlying structure. But examples from Serbo-Croatian in (12) show that vocative epithets do not present the only structural type of disapprobation expressions.

(12) a. Idiot! Serbo-Croatian
b. Jesi video šta sam uradio?! Budala.
AUX.2Sg see what AUX.1Sg done fool.NOM
'Did you see what I just did?! A fool.'
c. * Bobane$_i$, idiot$_i$!
Boban.VOC idiot.NOM
d. Jesi video šta sam uradio?! * Bobane, budala.
AUX.2Sg see what AUX.1Sg done Boban.VOC fool.NOM
e. Jesi video šta sam uradio?! * Ti/ja/on budala.
AUX.2Sg see what AUX.1Sg done you/me/he fool.NOM

In these examples, the noun with the expressive potential (*budala*, *idiot*) is used in nominative. Observe that no coreferential name in vocative, and not pronoun, including *ti* (you), can be added to this construction (12c-d) These expressions are used with a similar kind of expressive power. However, they do have a strong dimension of propositionality. Speakers agree that these expressions present some kind of title-like conclusion over a situation, and that they involve elision of a subject and a copular linker. This is confirmed by the examples in (13), which present common alternatives for those in (12), with approximately, or exactly, the same expressive potential, and even taking the same intonation pattern for the component bearing the expressive potential.

(13) a. Dakle: idiot! Serbo-Croatian
ergo: idiot
approximately 'Based on the context I conclude you're an idiot.'
b. Jesi video šta je uradio?! Pa on je budala.
AUX see what AUX done well he is fool.NOM
'Did you see what he just did?! Well, he's a fool.'

The speakers' intuitions about this construction are shared by the consulted Russian speakers for the examples of the type from P&R, as in (14).

(14) a. Idiot! Spoken Russian
idiot.MUSC
b. Idiotka!
idiot.FEM

Observe that, as (13b) shows, this construction can be used for any person, suggesting that something has been elided, i.e. that at least a *pro* is involved (which P&R indeed admit for the Russian example). It is impossible that small clauses without any access to the discourse involve a *pro*, since *pro* needs to be

bound from the discourse. All this shows that this second type of disapprobation expressions have more structure than a bare small clause, and that part of this structure is related to the discourse. They underlyingly present predicative constructions, with the subject and linker dropped due to a high topicality. This means that they do have a propositional status, and hence fail the requirements in (2a and b). It also makes them less interesting in respect of the questions discussed by P&R and in this paper, and therefore they will not be more thoroughly analyzed.

## 5 The status of expressive meanings

The analysis that I argue for in section **Fout! Verwijzingsbron niet gevonden.** above, that the English type of disapprobation expressions (P&R's disapprobation ESCs), just like the Serbo-Croatian ones, are in fact epithets in the vocative case, leaves the formal account of these constructions to the linguists dealing with vocatives and epithets. Whatever the explanation for why vocatives do not get embedded, and why they do not embed complex structures, that is also why disapprobation expressions in English and the relevant type in Serbo-Croatian do not do this. What is left is to decide whether disapprobation expressions of the discussed types bear regular semantic content, or do they derive a special, expressive type.

P&R commit to the latter view. They propose the following formal analysis for their view of disapprobation ESCs. Certain predicates, those with a strong expressive potential like *fool* or *idiot*, may be of the type <*a, E*> (in addition to their regular type, usually <*e, t*>), where *a* stands for any standard basic type element (i.e. <*e*> or <*t*>). *E* presents a novel type, representing the set of expressions whose value is in the domain of expressivity, and thus fundamentally different from the meanings of 'regular' linguistic expressions. In order to account for the fact that their disapprobation ESCs never take part in composing bigger structures, they stipulate that E is never the input for any complex type. This theory is nothing more than a description of the facts, and a costly one too: it introduces not only a whole new type (<*E*>), but also a special rule for this type: that it never embeds. They explicitly say that it is extremely difficult to describe the actual meaning of these expressions, and that getting to know more about them may lead to providing a more explanatory story for the type of meaning stipulated as <*E*> On the other hand, this is also a statement: that expressivity presents a proper semantic type of content, and a basic one, which does not derive from and which does not derive the standard types.

The analysis argued for in this paper reduces the number of reasons for introducing the type <*E*>. For instance, the semantic and syntactic isolation of the disapprobation expressions observed does not have to be stipulated as a

property of the type <*E*>, but takes any explanation that is offered for the considerable degree of semantic and syntactic isolation of vocatives. And since vocatives can be used with functions other than expressivity, this isolation is not a consequence of expressivity, but rather of some property that relates to all vocatives. The question that emerges is: Can we do without the <*E*>, and is there really such thing as a special, expressive, type of meaning? At least for the observed cases, this question can be approached by analyzing three important components that these expressions involve: 1) the descriptions used, such as *fool* and *idiot*, and their markedness in the relevant discourse 2) the epithet nature of the expressions (they do not use a neutral description in referring) and 3) the vocative case, i.e. a direct addressing of the collocutor (which may coincide with the speaker), and the fatic function that it has. Hopefully, the insight into these three elements, and how they conspire in disapprobation expressions, will lead to a proper analysis of the observed expressive meanings, making it unnecessary to introduce an additional basic type into the system of semantic representation.

P&R only present and discuss the negatively connotated disapprobation expressions, and not their positive counterparts like in (15).

(15) a. You little angel!
b. Srce malo!
hearth little
lit. 'You little hearth!'
c. Bobane, legendo!
Boban.VOC legend
lit. 'Boban, you legend!'

These examples, as well as those like (6), with neutral words used in disapprobation expressions, in a context that gives them a strong negative connotation, indicate that only predicates with strongly polarized connotations, whether by their lexical meaning or by the context, can be used as epithets in the vocative epithet type of expressions.

Moreover, all of them are at the same time hyperbolic. Not only a person cannot literally be a little hearth like in (15b), but also (15c) can only be used if to say for Boban that he is a legend is much stronger than his real role in the actual situation. Similarly in English, even if said to someone commonly considered an idiot, an expression like *you idiot* never simply states this fact, but still pretends on saying something stronger than the actual state of affairs.

Hence, the predicate of this type of expressive construction must have an exaggerated positive or negative connotation. This makes a good ground for expressive interpretations – the actual meaning of the predicate does not completely fit the pragmatics, and this mismatch signals that a special pragmatic

interpretation should be called for. Note, however, that since the predicate is used to refer, and not to assert, i.e. in the derivation of an <e> type expression instead of a <t> type one, the exaggeration of the predicate and its pragmatic status do not make it a new type. The expression still successfully refers, mostly due to the second person feature of the vocative, and the expressive power of the description has effects only in the purely pragmatic computation.

The second component, the epithet nature of this predicate, presents another signal for pragmatics that the expression is aimed at expressing more than its literal compositional meaning. Normally, whenever a proper name is available for a referent, it will be preferred to any descriptive expression, since it normally most economically and most unambiguously establishes reference. Even when a proper name is not available, reference is normally established using a description that involves the most contextually obvious properties of the referent, or was recently used to refer to it. Epithets used in the observed expressions are none of the above: they are not proper names, they do not use a property that belongs to the referent (but a hyperbolized one), and they do not present a description that has recently been used to refer to the same referent. The fact that the description presented by the epithet is not used as a predicate of a proposition, but in a referential expression, relying on the person feature and perhaps pragmatics to establish reference, is another aspect that signals a high level of pragmatic activation in the relevant expression.

Finally, the vocative case of the expression does two things. One is to establish a direct reference to the collocutor without requiring the description used to really describe the referent. This brings a particular kind of reference failure: the referent is unambiguously determined, but it does not fit the description used. The other is that vocative, with its fatic function, also bears a special status: it is syntactically quite simple, it stands out of the syntactic structure of the sentence, and it can bear strong and independent prosodic features (it can be, and often is, shouted).

As a result, the expressive interpretation of the observed construction comes about. The vocative expression is first interpreted literally, with two main components: the description that comes as its denotation and the vocative component marking that the expression refers to the collocutor. A clash emerges: the description is not among the neutral ways to refer to the collocutor: it is an epithet. Moreover, the epithet, literally interpreted, is among the properties presupposed NOT to hold for the referent. This signals the expressive nature of the expression. Finally, the epithet denotes a property with a strong positive or negative connotation, and pragmatics interprets this as an expression of a strong positive or negative attitude of the speaker towards the collocutor in relation to the local context of the expression. The collocutor, i.e. the referent of the expression, as already noted, can coincide with the speaker.

Thus we arrive at the expressive meaning of the observed expressions, deriving them from the components recognized in the analysis. Semantically, the expression has just its literal compositional meaning (in terms of type theory, based on the two standard types, <e> and <t>), with the final status of a cased nominal expression. This compositional meaning is pragmatically marked: it causes a failure in the description of the referent and introduces a hyperbolized positive or negative connotation, leading to a pragmatic interpretation that involves the expressive component.

## 6 The more general consequences

Based on their analysis of the disapprobation expressions, P&R draw very general conclusions in the domain of syntactic structure, its acquisition and the relation between two paradigms, the one of expressive and non-expressive meanings and the one of the corresponding structural inventories. They conclude that an expressive meaning can only be realized by an expression if it has no structural projections related to the discourse (such as TP and CP), and that in order to linguistically realize expressive meanings, adults resort to the two-word constructions characteristic of the early stages of acquisition. While children at the two word utterance stage generate only non-propositional expressive meanings which lack any interaction with the discourse, adults generate both, bare predications with expressive meanings and proper propositions with a rich discourse-related functional domain (CP).

The analysis presented in this paper questions many aspects of these conclusions. For instance, the vocative epithet construction is unlikely to be the same as the two word utterance used by children, because the latter normally have an addressee independent of the subject of the small clause. However, the main argument, that the observed type of expressions has no structure and no semantic material related to the discourse, might still hold. Notice that vocatives, whatever function they are used for, not only lack any elements related to the CP, but also lack the DP-related material. Definite nominal expressions, apart from proper names, do not appear as vocatives. No matter how clearly the context suggests the definiteness of an expression, if it includes the definite article, it is very degraded in the vocative use.

(16) a. (Hey!) *The waiter!
b. (Hey!) *The tallest guy down there!

It could be that vocatives are bare predicates, that they cannot project any discourse-related structure, and that this is why they fail to embed, and why they do not like to embed other structures. If so, this property would add to their

expressive nature, but at the same time, it would represent an exceptional and curious domain in language: a domain with non-propositional semantics. This is a highly important question, and an intriguing topic of research. The nominative type of expressive constructions like in (11) - (13), which appears to involve ellipsis and underlyingly present full sentences, hints that apart from the vocative case, as a marked category, and possibly several other similar categories, all linguistic expressions are trapped in their discourse-dependence. Once we acquire the discourse-related projections, unless they are blocked by something like the vocative feature (which is here just a stipulation and requires a thorough analysis to possibly be derived from something), they project in every linguistic expression that we produce. They all either establish reference (*v*P, DP) or update the discourse with new information (CP), and even if non-propositional meanings are to be realized, they are realized through propositions with a special pragmatic status. The only possible exception, vocative, indeed does not contribute new information to the discourse, and although it does establish reference, it is established by pragmatic means (contextually, by looking at the addressee, pointing etc.), while the linguistic expression involved seems to be at the level of a bare predicate.

## 7 List of references

Lebeaux, David. (1988) *Language acquisition and the form of grammar*. PhD thesis, University of Massachusetts, Amherst.

Pesetsky, David & Esther Torrego. (2000) T-to-C movement: causes and consequences. In Michael Kenstowicz (ed.). *Ken Hale: A life in language* Cambridge. MA: MIT Press.

Potts, Christopher & Tom Roeper (2006) "The narrowing acquisition path: from declarative to expressive small clauses." In Ljiljana Progovac, Kate Paesani, Eugenia Casielles-Suárez, and Ellen Barton (eds.) *The syntax of nonsententials: multi-disciplinary perspectives*. Amsterdam: John Benjamins.

Radford, Andrew. (1990) *Syntactic theory and the acquisition of English syntax*. Oxford: Blackwell.

Boban Arsenijević
University of Amsterdam
Department of Theoretical Linguistics
Spuistraat 210
1012VT Amsterdam
The Netherlands
Email: b.arsenijevic@uva.nl

# How to submit papers?

**Editors**: Miriam van Staden & Hedde Zeijlstra
**Contact**: H.H.Zeijlstra@uva.nl

**Publication form**: Web publication only in pdf-downloadables on the ACLC website. Links to DARE and (where applicable) group internal websites. Annual 'Volumes' consist of multiple 'Issues' that are filled incrementally as papers come in.

**Contents**: The ACLC Working Papers is first of all an informal platform to present new ideas and work in progress to fellow researchers inside and outside the ACLC, but it may also be used to pre-publish papers that have been submitted to (international) journals. Papers may take the form of a complete manuscript but also of a conference paper, discussion paper, report of an experiment, etc. *Authors retain all copyrights to their work* and are encouraged to publish (parts of) their paper in international journals later.

**Contributors**: We invite papers from ACLC members. External collaborators may be co-authors only. In the case of special issues external contributors may be accepted (e.g., in conference proceedings). We encourage especially PhD students to use this platform to share initial results and gain experience in writing academic papers.

**Instructions for authors**: Contributors may assume from their audience a fair amount of knowledge of the issues addressed in their papers. If peers from the relevant research group are familiar with a particular linguistic model, then authors do not have to go into detail in this respect. In this way, we hope to have concise contributions that delve right into the issues at hand.

Please find as a separate Word-document the 'style sheet' for the Working Papers. It contains various user-defined styles that authors may wish to make use of, e.g. for the formatting of examples. Note that tables and figures are put IN PLACE. Papers are submitted digitally in word-compatible format AND either as a pdf or as a hard copy. Any fancy artwork, special fonts, inserted video material, etc. is accepted so long as it can be incorporated into a pdf file. Publications in English and Dutch are accepted. Note carefully that all submitted papers must have been read by a native speaker. Please consult with the editors if you wish to deviate from these instructions.

**Reviewing process**: The editors in consultation with the director of the ACLC accept papers after a report from one or two referees. These referees will

typically be selected from the members on the editorial board, but the editors may decide that in particular cases experts from outside need to be consulted. Reports are requested only for papers that are rejected or accepted under condition of serious revision. The aim is to be quick but careful.

# Editorial Board

**Miriam van Staden**
**Hedde Zeijlstra**

## Editors in Chief

Miriam van Staden
Hedde Zeijlstra

## Board of editors

Enoch Aboh
Umberto Ansaldo
Nivja de Jong
Evelien Keizer
Harry Perridon
Rob van Son

## Contact

Dr H. Zeijlstra
Department of Dutch Linguistics
University of Amsterdam
Spuistraat 134
1012 VB  AMSTERDAM

H.H.Zeijlstra@uva.nl