Abstract. Few studies have investigated referential null subjects in present-day non-standard varieties of German. This paper explores the distribution of 4,000 null and overt referential subjects in finite clauses in a new corpus of spoken Swabian, a North Alemannic dialect spoken in Southwest Germany. The paper describes Swabian subject use as it emerges from the corpus data with regard to morphological form, syntactic distribution, reference and informational status. Swabian, unlike Standard German, allows and in some cases prefers clause-internal referential subjects to be null, whilst in other contexts, referential null subjects are ruled out. There is a robust second person null-subject property, with 61% of 2nd singular subject pronouns being dropped clause-internally, i.e. from a position following the finite verb in main clauses or following a complementiser in embedded clauses. 1st singular and 3rd singular neuter subjects are dropped clause-internally as well, but at much lower frequencies and only in very specific phonological and lexical constellations, not previously noted in the literature. The empirical findings are compared with an earlier unquantified study of Swabian (Haag-Merz 1996) and discussed in light of other work on present-day Germanic vernaculars and previous accounts of partial null subject languages.

1. Introduction

It is a well known fact that the present-day Germanic standard languages do not generally allow null referential subjects, and over the past three decades various explanations have been adduced to account for this: lack of strength in the verbal agreement paradigm to identify null subjects (e.g. Hyams 1986; Rohrbacher 1994), lack of morphological uniformity in the verbal paradigm to license null subjects (e.g. Jaeggli & Safir 1989), lack of Chinese-style topic identification of the null subject via a discourse topic (e.g. Huang 1984), or general incompatibility of V2 with referential null subjects (e.g. Jaeggli & Safir 1989).

At the same time, a growing number of empirical studies show that contrary to expectation, non-standard Germanic dialects do allow referential null subjects, and these are by no means restricted to sentence-initial topic-drop contexts, but occur clause-internally (cf. e.g. Bayer 1983/1984, Weiβ 1998 and Fuß 2005 for Bavarian; Cooper & Engdahl 1989 and Werner 1999 for Zurich German; Haag-Merz 1996 for Swabian; de Haan 1994 and Hoekstra 1997 for Frisian). None of the modern Germanic dialects studied to date appear to allow referential null subjects across the board (pro-drop), but they only do so for certain number(s) and grammatical person(s), usually for second person and sometimes for first person too (partial pro-drop). This may be linked to recoverability of subject phi-features via corresponding distinct verb inflections that exist in these Germanic vernaculars. For many varieties
however, empirical data are lacking concerning the precise syntactic distribution of referential null subjects. Moreover, hardly any quantified data exist that allow us to investigate the frequency of null vs. overt referential subjects, the frequency of null subjects in certain contexts, or the frequency of null subjects with a particular interpretation/person reference.

The present paper is a step towards filling this empirical gap, as it for the first time provides quantified data from a corpus of spoken Swabian. Swabian is a North Alemannic German dialect spoken by ca 2 million speakers, who, save for the very old, all are bidialectal due to the Standard German schooling system and the media. Since Swabian is often regarded as uncouth/non-academic, educated speakers typically avoid speaking it outside all-Swabian contexts. Spontaneous oral data were collected from three 30-year-old female speakers with similar backgrounds (same variety of rural Highland Swabian, monolingual Swabian upbringing, exposure to Standard German through school), as well as from three elderly female Swabian speakers, in one-to-one dialogue with a linguist native speaker of the same variety to avoid accommodation. The resulting corpus contains 30,700 words and 4,000 instances of subjects.

Focusing on subject pronouns and null subjects, the paper describes the pronominal systems of the speakers as they emerge from the corpus data with regard to morphological form, syntactic distribution, reference and informational status. We will see that Swabian, unlike Standard German, allows and in some cases prefers clause-internal referential subjects to be null, whilst in other contexts, referential null subjects are ruled out. In particular, the corpus data indicate that Swabian has a robust second person null-subject property, with 61% of 2nd singular subject pronouns being dropped from inside the clause, i.e. from a position following the finite verb in main clauses or following a complementiser in embedded clauses (as illustrated in (1)).

(1) [Context: A notoriously talkative acquaintance is ringing the doorbell …]

\[
\text{wenn } \emptyset \text{ ez dui rei losch, kommsch } \emptyset \text{ zo nix mee if } \emptyset \text{ now her in let-2SG come-2SG } \emptyset \text{ to nothing more}
\]

‘If (you) let her in now, (you) won’t get anything done any more.’

Apart from 2nd singular, Swabian also allows certain other types of clause-internal null subjects (1st singular and 3rd singular neuter); these however occur at much lower frequencies (1% and 2.5%, respectively) and only in very specific phonological and lexical constellations, not previously noted in the literature. On the other hand, clause-internal subjects of other persons and numbers (1st plural, 2nd plural, 3rd singular masculine, 3rd singular feminine, 3rd plural, 3rd polite, 3rd generic) are never null in Swabian. The empirical findings are compared with an earlier unquantified study of Swabian (Haag-Merz 1996) and
discussed in light of other work on present-day Germanic vernaculars and previous accounts of (partial) null subject.

This paper is set up in the following manner. Section 2 briefly summarises existing studies of null subjects in Swabian. Section 3 describes the informants, data collection, transcription, coding and method of analysis of the present study. An overview of the results is given in section 4, whilst the more detailed quantitative and qualitative findings concerning Swabian 2nd singular null subjects, 1st singular null subjects and 3rd singular null subjects in the corpus are described in sections 5, 6 and 7. Here, subsections are devoted to an investigation of the precise syntactic environments that clause-internal referential null subjects occur in, and to the potential influence of discourse-pragmatic, lexical and phonological factors. Section 8 contains a concluding discussion.

2. Previous studies

A number of dialect studies and diachronic work on older varieties of German, e.g. Old High German, mention Swabian as one of the modern varieties of Germanic that allow second person singular subjects to be dropped (e.g. Cooper & Engdahl 1989:43; Fuß 2005:219; Axel 2007:321; Rosenkvist 2009:164). However, such mention is typically no more than a brief footnote. As far as I am aware, there is only one study that has investigated null and overt subjects in Swabian to date. Haag-Merz, in the first part of her 1996 dissertation, introspectively considers subject and object pronouns in her own native dialect of urban Lowland Swabian (Stuttgart/Remstal). The second part of her dissertation is a case study of the acquisition of pronouns in a Swabian-speaking child. Haag-Merz characterises adult Swabian as partial pro-drop, where the omission of subjects is restricted to first person singular (1SG) and second person singular (2SG). She argues that 2SG subject drop is a syntactic option and widespread in Swabian, whereas 1SG subject drop is phonologically conditioned and more restricted (1996: Ch. 3). Null subjects for other persons and numbers are said not to exist. Haag-Merz (1996) constitutes an important work, but as it is example based and lacks quantitative data, it does not investigate the distributional frequencies of null subjects. Haag-Merz’s work will be discussed in light of the empirical findings to be reported on.

3. Informants, data collection and method

The empirical data of the present study come from a new corpus of speech recordings of six native speakers of Swabian matched for age, background and area of origin. Spontaneous oral data (30,700 words) were collected between 1994 and 2000 from three younger female
speakers with similar backgrounds (same variety of rural Highland Swabian, monolingual Swabian upbringing, exposure to Standard German (Hochdeutsch ‘High German’) through school), as well as from three elderly female speakers, in one-to-one dialogue with a linguist native speaker of the same variety to avoid accommodation. Informants from two different generations were included to see whether there might be any age-related differences in their Swabian.

The 3 younger informants were all between 25 and 35 years of age at the time of the recording. They were born in and had grown up in very small towns in a rural area with a radius of 12 miles at the south-eastern edge of the Swabian Highlands in the vicinity of the city of Ulm. Their families including their parents were from the area, had lived there all their lives and spoke Swabian at home. The informants had spent their early childhood mostly at home with their mother, plus some local part-time nursery care. Regular exposure to High German started at the age of 6 or 7 at the local state school, where Standard German was taught through reading, writing, some teacher talk, and the media, whereas conversation between pupils and most teachers, between pupils and at home remained Swabian. Having completed grammar school at 19, they left home for university and spent much of the following years abroad and in university environments in southwest Germany. Nowadays they use Swabian mainly with relatives and friends back ‘home’. In this context it is noteworthy how strong their native dialect remains.

The 3 older informants were all between 60 and 70 years of age at the time of recording. They came from the same area as the younger informants, had grown up there and had lived there all their lives. In contrast to the younger informants, they only had primary and some secondary-level or vocational education. The older informants spoke Swabian on a daily basis in and outside the home, their active use of High German being limited to occasional conversations with non-Swabian outsiders.

I audio-recorded 45 minutes of spontaneous speech for each speaker in conversation with me, myself also a native speaker of Swabian, with the same background as the younger informants. Using a High German-speaking interviewer would not have worked, as the speakers would then have switched to an approximation of High German themselves. Recording would start after an hour or so of warming-up conversation. In the recordings, the speakers spontaneously discuss current events, describe procedures, e.g. passport applications or biscuit making, and narrate past events such as holiday activities. Conversation in Swabian felt natural, was easy going and Swabian was kept up continuously, with only occasional switches into High German when the speaker was quoting a non-Swabian speaker. The speakers were unaware of the purpose of the recording. I led them to believe that I was investigating differences between Swabian and High German pronunciation.
I transcribed the data in an adapted CHAT format (MacWhinney 2000). Swabian is primarily an oral language, without any standardised transliteration system for spoken Swabian. The general aim was to transcribe words the way they were pronounced. As my prime objective was morphosyntactic, I did not carry out any (time-consuming) phonetic transcriptions. I did want to capture the spoken language without introducing potentially misleading Standard German transliteration conventions, but I also wanted the transcripts to be decipherable for readers familiar with German. Thus, a modified orthography was used. This meant that I followed the rules of German orthography where the Swabian rendering matched that of Standard German. In most cases, Swabian differs from Standard German and here a broad phonemic transcription was used, detailed enough to allow for morphological analysis, but not necessarily for phonological analysis.

Swabian contains many proclitics, enclitics and clitic clusters and I transcribed these together with their host word as one orthographic word, but marked morpheme boundaries in the case of pronominal subject clitics. Had the utterances been Standard German, many of the Swabian clitics (e.g. proclitic and enclitic pronouns, articles and particles) would have been translated as non-clitics and thus been orthographic words in their own right. As a result, the 30,700 word count would have been substantially higher if the utterances had been transcribed according to Standard German conventions. Stress was also marked so that searches for e.g. stressed and unstressed subject pronouns would be possible. In keeping with CHAT conventions, exclamation marks inside square brackets […]! indicate elements that carry focal stress (MacWhinney 2000).

I then manually classified and coded the transcribed utterances. Roughly 4,000 subjects in finite clauses were found in the material. These subjects were classified according to person (1st, 2nd, 3rd), number (singular, plural) and grammatical gender (masculine, feminine, neuter) as well as according to whether they were lexical, pronominal or null. Pronominal subjects were also classified according to morphological form and whether they were stressed or not.

1 Some written Swabian texts do exist in the public domain (mainly song lyrics, humorous newspaper columns, poetry), with a variety of idiosyncratic transliterations. However, the lack of any standardised transliteration conventions effectively prohibits literacy in Swabian, and reading and writing are not taught.

2 Portions of the recordings were transcribed twice, independently, using different types of transcribing equipment. One transcription was made right after collection, a second version after a considerable time interval. Each version was checked against the tape, at different speeds and noise levels. Recording quality was generally good. I then compared the two versions of the transcripts. Any resulting differences were resolved by checking the transcripts again against the tape. I therefore assume that the transcripts are reliable, even though no additional independent transcriber was available.
Linear position of the subject (clause-initial vs. clause-internal) was coded as well. Clause-initial \(_V_{FIN}\) refers to subjects in the first position of a main clause, thus preceding the finite verb (verb second). ‘Clause-initial’ is thus the position where omission of the subject is typically referred to as ‘topic drop’ or ‘pronoun zap’. I did investigate such cases, but they will not be the focus of the present paper. ‘Clause-internal’ position was broken down into postverbal subjects in main clauses (following the finite verb, \(V_{FIN}\)) and post-complementiser subjects in embedded clauses (C\(_\_\)). These two clause-internal positions are positions where subject omission is commonplace in pro-drop languages (such as Italian and Spanish) but said to be ungrammatical in non-pro-drop languages (such as Standard German and the standard Germanic languages). These cases will be the main topic of the present paper. Elliptical subjects in coordinated clauses (also known as conjunction reductions, e.g. *she is German and _ lives in Sweden*) were coded as a separate category. There were also some unclassifiable subjects.\(^3\) I concentrate on the linear order of constituents and abstract away from analytical questions concerning the structural account of this linear order, as none of the findings to be reported on hinges on specific syntactic analyses.

4. Overall results

The following tables provide an overview of the overall distribution of the 4002 subjects in the Swabian corpus. The majority of subjects is pronominal (79%), followed by lexical subjects and null subjects, as shown in Table 1. This predominance of pronominal subjects is typical of many corpora of informal speech. Null subjects occur at a rate of 10.5%, which is noteworthy for a Germanic language. These 10.5% include both clause-initial and clause-internal subjects and will be investigated further below.

**Table 1.** Form of subjects overall

<table>
<thead>
<tr>
<th>Lexical subject</th>
<th>Pronominal subject</th>
<th>Null subject</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>424</td>
<td>3159</td>
<td>419</td>
<td>4002</td>
</tr>
<tr>
<td>10.6%</td>
<td>79.0%</td>
<td>10.5%</td>
<td>100%</td>
</tr>
</tbody>
</table>

\(^3\) Unclassifiable subjects were of two types: (i) utterance fragments where the linear position of the subject could not be determined, and (ii) relative pronouns that introduce an embedded relative clause and simultaneously function as its subject (e.g. bolded *der ‘who/that’* in the following example: *da had [der!], der den [klein!] bus gfahren had … then has HE that-sg.masc the minibus driven has ‘then the guy who drove the minibus …’*). Without a separate complementiser, the position of such subject relative pronouns in relation to C could not be determined, and these cases were counted as unclassifiable.
As Table 2 shows, most subjects (lexical, pronominal and null) occur in main clauses, with an even distribution between clause-initial preverbal positions (41%, \textsubscript{VFIN}) and postverbal positions (42%, \textsubscript{VFIN\_}). Only a minority of subjects occur in embedded clauses, after the complementiser (14%, \textsubscript{C\_}).\textsuperscript{4}

Concerning number and person reference, there is an uneven distribution in the Swabian corpus (Table 3), with most of the subjects being 1st singular (26%) or 3rd singular (45%, all genders combined). There are substantially fewer 2nd singular subjects (9%), 1st plural subjects (5%), 3rd generic subjects (5%) and 3rd plural subjects (10%), and only very few 2nd plural and 3rd polite subjects. One should perhaps not read too much into this uneven distribution, as genre, modality and type of discourse are known to heavily influence the frequencies with which subjects of a particular number and person occur in a corpus.

Table 2. Position of subjects overall

<table>
<thead>
<tr>
<th>Clause-initial: \textsubscript{VFIN}</th>
<th>Clause-internal: \textsubscript{VFIN_}</th>
<th>Clause-internal: \textsubscript{C_}</th>
<th>Null coordinated</th>
<th>Unclassifiable</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1645</td>
<td>1689</td>
<td>561</td>
<td>68</td>
<td>39</td>
<td>4002</td>
</tr>
<tr>
<td>41.1%</td>
<td>42.2%</td>
<td>14.0%</td>
<td>1.7%</td>
<td>1.0%</td>
<td>100%</td>
</tr>
</tbody>
</table>

As Table 2 shows, most subjects (lexical, pronominal and null) occur in main clauses, with an even distribution between clause-initial preverbal positions (41%, \textsubscript{VFIN}) and postverbal positions (42%, \textsubscript{VFIN\_}). Only a minority of subjects occur in embedded clauses, after the complementiser (14%, \textsubscript{C\_}).\textsuperscript{4}

Concerning number and person reference, there is an uneven distribution in the Swabian corpus (Table 3), with most of the subjects being 1st singular (26%) or 3rd singular (45%, all genders combined). There are substantially fewer 2nd singular subjects (9%), 1st plural subjects (5%), 3rd generic subjects (5%) and 3rd plural subjects (10%), and only very few 2nd plural and 3rd polite subjects. One should perhaps not read too much into this uneven distribution, as genre, modality and type of discourse are known to heavily influence the frequencies with which subjects of a particular number and person occur in a corpus.

Table 3. Distribution of subjects overall according to person, number and gender

<table>
<thead>
<tr>
<th></th>
<th>1SG</th>
<th>1PL</th>
<th>2SG</th>
<th>2PL</th>
<th>3SG- MASC</th>
<th>3SG- FEM</th>
<th>3SG- NEU</th>
<th>3- POLITE</th>
<th>3- GENERIC</th>
<th>3PL</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1037</td>
<td>199</td>
<td>368</td>
<td>10</td>
<td>328</td>
<td>368</td>
<td>1084</td>
<td>15</td>
<td>192</td>
<td>401</td>
<td>4002</td>
</tr>
<tr>
<td></td>
<td>25.9%</td>
<td>5.0%</td>
<td>9.2%</td>
<td>0.3%</td>
<td>8.2%</td>
<td>9.2%</td>
<td>27.1%</td>
<td>0.4%</td>
<td>4.8%</td>
<td>10.0%</td>
<td>100%</td>
</tr>
</tbody>
</table>

In what follows, I will look at the distribution of null and overt subjects for each person, number and gender separately and in relation to the context they occur in. Overall, there are 6.8% clause-internal null subjects (Table 4 below) and 5.6% clause-initial null subjects (Table I in fn. 7). Null subjects in coordinations (1.7%) and unclassifiable subjects (1.0%) which are included in Table 2 for the sake of completeness will be disregarded in the remainder of this paper.

The speakers produced the following forms of pronominal subjects clause-internally (2). By and large, they corroborate the pronominal system Haag-Merz (1996:91) has proposed for Swabian on the basis of her own intuitions, though the exact phonological forms differ slightly depending on region and rural/urban aspect.

\textsuperscript{4} This asymmetry is perhaps not that surprising, since main clauses are nearly six times as frequent in the corpus (N = 3387) as subordinate clauses (N = 594).
When null and overt subjects are broken down by individual person, number and gender, a striking picture emerges (Table 4, Figure 1).

Table 4. Clause-internal null subjects

<table>
<thead>
<tr>
<th></th>
<th>1SG</th>
<th>1PL</th>
<th>2SG</th>
<th>2PL</th>
<th>3SG-MASC</th>
<th>3SG-FEM</th>
<th>3SG-NEU</th>
<th>3-POLITE</th>
<th>3 GENERIC</th>
<th>3PL</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>13</td>
<td>0</td>
<td>219</td>
<td>0</td>
<td>0/357</td>
<td>0/357</td>
<td>0/357</td>
<td>0/191</td>
<td>0/16</td>
<td>2/184</td>
<td>257/3803</td>
</tr>
<tr>
<td>1PL</td>
<td>995</td>
<td>0</td>
<td>0/185</td>
<td>0</td>
<td>0/357</td>
<td>0/357</td>
<td>0/357</td>
<td>0/191</td>
<td>0/16</td>
<td>0/375</td>
<td>257/3803</td>
</tr>
<tr>
<td>2SG</td>
<td>0</td>
<td>0</td>
<td>6.8%</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5.6%</td>
<td>6.8%</td>
</tr>
<tr>
<td>2PL</td>
<td>0</td>
<td>0</td>
<td>6.8%</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5.6%</td>
<td>6.8%</td>
</tr>
<tr>
<td>3SG-MASC</td>
<td>0</td>
<td>0</td>
<td>6.8%</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5.6%</td>
<td>6.8%</td>
</tr>
<tr>
<td>3SG-FEM</td>
<td>0</td>
<td>0</td>
<td>6.8%</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5.6%</td>
<td>6.8%</td>
</tr>
<tr>
<td>3SG-NEU</td>
<td>0</td>
<td>0</td>
<td>6.8%</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5.6%</td>
<td>6.8%</td>
</tr>
<tr>
<td>3-POLITE</td>
<td>0</td>
<td>0</td>
<td>6.8%</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5.6%</td>
<td>6.8%</td>
</tr>
<tr>
<td>3 GENERIC</td>
<td>0</td>
<td>0</td>
<td>6.8%</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5.6%</td>
<td>6.8%</td>
</tr>
<tr>
<td>3PL</td>
<td>0</td>
<td>0</td>
<td>6.8%</td>
<td>0</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5.6%</td>
<td>6.8%</td>
</tr>
</tbody>
</table>

5 Similarly to Standard German, Swabian has a polite form sia/se ‘you’ that is used to address strangers and persons of higher standing. The form is homophonous with and historically derived from the 3rd plural personal pronoun sia/se ‘they’, and like this 3PL pronoun, polite sia/se combines with verbs inflected for 3rd plural. In the corpus at hand, the polite form occurs infrequently, since the interviewer and the informants are on familiar terms with each other, typically addressing each other with familiar du ‘you’.

6 Similarly to Standard German, Swabian has a generic subject pronoun ma/mr ‘one’ that is historically derived from the noun for ‘man’.

7 Clause-initial null subjects (‘topic drop’), null subjects in coordinate structures and unclassifiable subjects are not included in the counts in Table 4 and will not be discussed any further in the paper. For the sake of completeness, I provide figures for clause-initial null subjects by person, number and gender in Table 1 below. 1PL, 2PL, 3-POL and 3-GEN subjects are never null, though this could be an artifact of the low totals for these subject categories in the corpus in general. For 3SG-MASC, 3SG-FEM and 3PL, we find a few clause-initial null subjects. 1SG, 2SG and 3SG-NEU subjects are dropped from clause-initial position at more substantial rates (5%, 7%, 10%). Interestingly, 1SG, 2SG and 3SG-NEU are precisely those categories that can also be dropped clause-internally. Percentages are however quite different for clause-initial vs. clause-internal position (see Table 4).
The null subject phenomenon is selective and only occurs for part of the verbal paradigm: Most notably, 61.3% (219/357) of the 2nd singular subject pronouns are dropped from inside the clause, i.e. from a position following the finite verb or a complementiser, as illustrated in examples (3)–(4) and (5) respectively. Ø indicates the position of the null subject, based on where an overt pronominal subject would (and could) occur in the utterance.

(3) [About feeling some pain after sitting still for stretches of time]  
  ah, [des!] *hosch* Ø jetzt gmerkt.  
  ‘Oh, (you)’ve noticed that.’

(4) [How to make pastry for a plum cake]  
  doo *duasch* Ø dann au a ganz [oi!] nei, ond ver[kneeda!]  
  there do-2SG _ then also a whole egg in and together-knead  
  ‘Well then (you) also add one egg and knead it together.’

(5) [About watching your step on an cliff path]  
  weil *wenn* Ø daa runder kucksch, des isch [total!] [tief!]  
  because if _ there down look-2SG it is totally deep  
  ‘Because if (you) look down there, it’s awfully deep!’

The Swabian corpus also contains clause-internal null subjects with 1st singular reference (1.3%) and 3rd singular neuter reference (2.5%), exemplified in (6) and (7), though these are much less frequent than the 2nd singular null subjects (61.3%).

---

Figure 1. Referential null subjects broken down by person, number and gender

© 2013 The Editorial Board of Studia Linguistica
10 Ute Bohnacker

(6) [About some homeopathy information sheet inside a parcel]
vielleicht kommt des in dem [zeddele!] doo drenna, wo Ø
maybe comes it in the sheet there inside that _
dr neidoa han.
you-DAT in-put have.1SG
‘Maybe it says so on the sheet that (I) put in there for you.’

(7) [About adding only one lump of crystallized sugar to her drink]
nee, i brobier-s erschmol mid [oim!] … daß Ø dann ned
no I try.1SG-it first with one … that _ then not
zu [süß!] wirt.
too sweet gets.3SG
‘No, I’ll try it with ONE lump first, so that (it) won’t get too sweet.’

By contrast, subjects of other persons and numbers (i.e. 1st plural, 2nd plural, 3rd singular masculine, 3rd singular feminine, 3rd plural, 3rd polite, 3rd generic) are never dropped clause-internally in the Swabian corpus (Table 4, Figure 1). This indicates that Swabian allows null subjects only in very restricted environments. Different conditions for null subjects depending on person may be operative in the language at the same time, and the following sections investigate these in more detail.

We will start with 2nd singular, as these are the most common null subjects.

5. Results: Second person singular null subjects

In the aggregated data from all six informants, there are 368 contexts for a second person singular (2SG) subject to occur, including 357 clause-internal contexts, and in the majority of these, the subject is null (61%, 219/357). The proportion of null 2SG subjects ranges between 50% and 78% for the individual speakers, without any discernible age-related differences.

5.1. An asymmetry between main clauses and embedded clauses?

Clause-internal 2SG null subjects occur postverbally in main clauses after any type of finite verb (auxiliary, copula, thematic verb) and post-complementiser in embedded clauses after various complementisers (dab, ‘that’, wenn ‘if/when’, weil ‘because’, wo ‘when’).\(^8\) Interestingly, postverbal 2SG subjects in main clauses are mostly null (71% null (210/295)), whereas 2SG subjects in embedded clauses are mostly overt: only 14% (9/

\(^8\) Present-day Swabian does not inflect its complementisers, and neither did older varieties of the language as far as I am aware (Haag-Merz 2006:157–158). To confirm the complete lack of inflecting complementisers in Swabian, I went through all 432 instances of complementisers introducing an embedded clause in the corpus. None carried any inflection.
Null subjects in Swabian

65) are null after a complementiser. Null forms are illustrated by (1), (3)–(5) above and the examples in (8)–(11).

(8) vielleicht merksch Ø plötzlich [nix!] mee.
maybe notice-2SG _ suddenly nothing more
‘Maybe (you) suddenly won’t feel anything any more.’

(9) [About going for an evening walk in Iceland in July]
isch ja dann immr [hell!], da kannsch Ø ja dann
is well then always light there can-2SG _ well then
immr no.
always still
‘Well, it’s always light then. So (you) can still go (for a walk) then.’

(10) [Discussing alternatives to watching television]
i mein, was machsch-Ø-n da au, wenn Ø kein, wenn
I mean what do-2SG-_-then there though if _ no if
Ø kein [fernseher!] und sonsch no ebbes hosch.
_ no television and else still anything have-2SG
‘I mean, what else can (you) do though if (you) haven’t got a
television or anything.’

(11) [About spending the night in a heated hut whilst on a horse-trek]
abr daß Ø de hald [auf!]wärmen konndesch un
but that _ 2SG.REFLEX well up-warm could-2SG and
widdr was # was [essn!] konndesch.
again what what eat could-2SG
‘…well, so that (you) could get warm and have something to eat
again.’

Each of the six informants regularly drops 2SG subjects from postverbal position in main clauses, but only three informants drop 2SG subjects after a complementiser in embedded clauses. Does this mean that the other three informants do not allow post-complementiser null subjects? It may be premature to draw such a conclusion from the corpus data, since the number of contexts for complementiser + 2SG is not particularly high.

Still, the asymmetry between 2SG subjects after a finite verb in a main clause and after a complementiser in an embedded clause is noteworthy, because several authors have recently made claims about the distribution of referential null subjects in main clauses vis-à-vis embedded clauses. For instance, Rosenkvist (2009:171) maintains that “in all of the modern Germanic vernaculars, RefNSs [referential null subjects] are possible in both main clauses and in embedded clauses, and there do not seem to be

9 All 65 embedded clauses with (null or overt) 2SG subjects have verb-end word order. Such clause-final placement of the finite verb is a hallmark of subordinate clauses.
10 One of the older and two of the younger informants drop post-complementiser 2SG subjects occasionally. Postverbally, they drop 2SG subjects much more frequently.
any differences in frequency related to clause type – again, this is in sharp contrast to the distribution of RefNSs in Old Germanic” [my emphasis]. Rosenkvist’s claim concerning modern Germanic vernaculars explicitly includes Swabian but is not backed up with any frequency counts, whereas the corpus data presented here suggest that there is in fact a clear difference concerning the frequency of null 2SG subjects for different clause types in Swabian.

Other researchers working on modern Germanic dialects closely related to Highland Swabian have indeed suggested that 2SG subjects are dropped more easily in main clauses than in embedded clauses. Haag-Merz (1996) argues this for urban Lowland Swabian based on her own native-speaker intuitions, and Werner (1999) suggests the same for Zurich German on the basis of speech samples and an acceptability judgment task: Werner’s 12 informants accepted 2SG null subjects in postfinite position in main clauses much more so than in post-complementiser position in embedded clauses (1999:45, 59–62). Even though these authors do not provide any quantification, Haag-Merz (1996:153–156) and Werner (1999:62) describe null 2SG subjects in main clauses as the more common, default option in Lowland Swabian and Zurich German, and null 2SG subjects in embedded clauses as a possible but rarer option. My corpus data point in the same direction: In spontaneous Swabian speech, 2SG subjects in main clauses are dropped much more frequently (71%) than in embedded clauses (14%).

Other authors have suggested that referential pro-drop in modern Germanic dialects can occur in embedded clauses only if the dialect has complementiser agreement, i.e. inflected complementisers that agree with the subject in person and number (e.g. Bayer 1983/1984; Hoekstra 1997; Weiß 1998:119–133, 2005; Fuß 2004:82; Axel 2007:321, 329; Axel & Weiß 2010:16–21, to appear).¹¹ These authors adduce data from Bavarian, West Flemish and Frisian, which have both inflecting complementisers and null subjects in embedded clauses, and contrast them with language varieties that neither have complementiser agreement nor null subjects in embedded clauses. Unlike Bavarian, West Flemish and Frisian however, Swabian does not have complementiser agreement, yet 2SG null subjects in embedded clauses occur in the corpus all the same, as we have seen. This means that complementiser agreement is unlikely to be a universal prerequisite for referential pro-drop.

¹¹ This phenomenon is considered complementiser agreement because clauses with subject-agreement inflection on the finite verb and subject-agreement inflection on the complementiser may additionally contain an overt pronominal subject. This is illustrated by the example in (i) from Bavarian (Bayer 1983/84, his example (100)):

(i) wenn-st (du) kunnst
when-2SG you.2SG come-2SG
‘if/when you come’
5.2. A pragmatic explanation for 2SG null subjects?

The majority of the 2SG subjects in the Swabian corpus are null (61%), and one might therefore want to consider the null form as the default for subjects of this person and number. In this case there should be specific contexts where the null default is ‘overruled’ by an overt form for a particular reason, for instance some pragmatic reason. Here, it has been suggested that language varieties with null subjects as the default employ overt pronouns to signal contrastive focus on the subject, to signal new (rhetic) information, or to signal a change of topic. This means that once a subject has been established as a topic, it must be followed by a sequence of null subjects in both main and embedded clauses. An overt subject pronoun introduces a contrastive focus or a change of topic interpretation (e.g. Y. Huang 1995; Carminati 2002; Öztürk 2002; Frascarelli 2007; Cole 2009 for Italian, Spanish, Turkish and Chinese). Similar proposals have been voiced for modern Germanic as well (though not quite in so much detail). For instance, according to Haag-Merz (1996:153) focused subjects in Lowland Swabian must be overt but non-focused 2SG subjects are null, and Weiß (1998:128) makes a similar claim concerning Bavarian. Rosenkvist (2009) characterises the modern Germanic vernaculars as follows: “an overt subject (which could have been omitted) appears to trigger emphatic/contrastive interpretations in all of these language varieties” (Rosenkvist 2009:172), though he then qualifies this claim by saying that “there are not clear data concerning the pragmatic effects of overt subjects from all of the language varieties” (2009:172). To test these proposals for Swabian, I investigated the discourse-pragmatic contexts in which 2SG subjects are null versus overt.

Of the 357 clause-internal 2SG subjects in the corpus, 8 are stressed pronouns (\[\text{du!}\]). They all occur in contexts where the subject is in focus, as in (12) and (13).

(12) [A asks B whether she should give her a particular book, having forgotten that she has already sent B this book. B reminds A that it was A who sent it to her]

\begin{center}
B: des hosch [\text{du!}] doch [\text{mir!}] scho mool [kschigged!].
\end{center}

\begin{center}
it have-2SG you but me already once sent
\end{center}

‘But YOU’ve sent me this already.’

© 2013 The Editorial Board of Studia Linguistica
(13) [A has given B some keys and papers so that B can act on her behalf whilst A is away, but B wonders whether A might still need them herself before leaving]

B: also wenn [du!] jetzt ed abr no was viellei
doh so if you now not but still some maybe

[brauchsch!], woissch?

need-2SG know-2SG

‘But what if YOU still need some of this (before you leave), you know?’

A: noi, [i!] brauch jetz [nix!] mee.

no I need now nothing more

‘No, I don’t need ANY (of this) any more.’

When comparing the 8 occurrences of stressed 2SG pronouns ([du!] ‘YOU!’) with the much more frequent occurrences (130) of unstressed 2SG pronouns (du, da, de, d ‘you’) in the corpus, it turns out that none of the unstressed pronouns occur in contexts where the subject is focused; rather the subject is informationally unimportant (background, old information), just like in cases when a null 2SG subject is used. An example is given in (14).

(14) [A and B discussing the potential effects of herbal infusions]

B: des hab i damals ja monade [gmacht!] mid diesem that have I then well months done with this
dumma frauamandl.

stupid lady’s mantle

‘I DID that for months, drinking that stupid lady’s mantle (infusion).’

A: des [woiß-i!] guad, des woiß i [guad!].

that know-I well that know I well

‘I KNOW, I do know that.’

A: doo hosch Ø ja ksagt +...

then have-2SG _ well said

‘And then (you) said…’

A: [mei!] was hosch du den [tee!] tronka.

gee what have-2SG you the tea drunk

‘Gee, you really drank that tea a lot.’

In (14), A addresses and refers to B in the 2nd person but does not use an overt 2SG pronoun at first but a null subject (hosch + Ø). The ‘pragmatically superfluous’ overt 2SG pronoun comes later (hosch + du). A could also have used two null pronouns or two overt pronouns, and sometimes does so in similar dialogues. Null 2SG subjects thus appear to occur in free variation with unstressed 2SG pronominal subjects (du, da, de, d).
Consider also the following exchange (15). Speaker A is continually addressing B throughout, B is the 2SG subject and topic of all sentences, and there is no change of referent. A asks B whether she will be paid for an upcoming business trip by using a null subject (kriagsch + Ø), and since B does not supply the answer, A asks her once more, again employing a null subject (kriagsch + Ø). A then inquires when B will be leaving for her trip, using an overt unstressed du subject in her question (wann gehsch + du). However, there has been no change of referent, no change of topic, du is clearly old information in this context, and du is neither focused nor contrasted; rather, focal stress is on the wh element and the verb.

(15) [A and B discussing a business trip B will be going on]

A: kriagsch Ø au [zahlt]!?”
   get-2SG _ though paid
   ‘Are (you) getting paid though?’
B: bidde?
   ‘What?’
A: kriagsch Ø des zahlt?
   get-2SG _ it paid
   ‘Are (you) getting paid for it?’
B: ja.
   ‘Yes.’
A: ond [wann!] [gehsch!] du dann jetz?
   and when go-2SG you then now
   ‘And when are you leaving now again?’
B: am [sonntag]!
   ‘On Sunday.’

In the corpus data, there are many exchanges as in (14) and (15), where an utterance with a 2SG null subject is followed by an utterance with an unstressed overt subject. It is unlikely then that the 2SG subject has suddenly become overt because of a pragmatic reason, such as the need to reactivate an otherwise unclear referent, to change the topic, or to focus the subject. There are also many instances where 2SG unstressed overt subjects are followed by 2SG null subjects, where 2SG unstressed overt subjects are followed by overt subjects, and where null subjects are followed by null subjects, without any discernible differences in interpretation. To my mind, this suggests that 2SG null and unstressed overt subjects in Swabian are indeed in free variation, and overtness is not determined by discourse-pragmatic factors as may be the case for Italian, Spanish and Turkish for instance.

Going back to the aforementioned claims in the literature on modern Germanic varieties about the informational status of different 2SG subject realisations (e.g. Rosenkvist 2009), the Swabian corpus data indicate that focus/contrast on the subject is not expressed by
using an overt subject pronoun instead of a default null form, but rather by using a stressed overt subject pronoun instead of an unstressed or null form.

5.3. **Phonological and lexical restriction on 2SG subjects?**

To explore the optionality of clause-internal 2SG null and unstressed overt subjects further, I looked at the phonological environments of the two in the corpus, but could not detect any differences. There were even a number of utterances where the speaker produced near-identical utterances – save for the null/overt pronoun.

I also considered whether 2SG null subjects might be restricted to certain kinds of verbs whilst overt subjects would occur with other verbs. This did not turn out to be the case, as speakers produced both null and overt subjects with the very same verbs, as illustrated in the a. and b. pairs in (16)–(17) produced by the same speaker; recall also the minimal pair (*hosch* + Ø ‘you have’ and *hosch* + *du* ‘you have’) from example (14).

(16) a. un [dann!] **mächsch du** des moodl [naß!].
   and then make-2SG you the mould wet
   ‘And THEN you make the mould WET.’

   b. do **mächsch Ø** im brinzib en rechd [oifacha!] doig.
   then make-2SG _ in principle a quite simple dough
   ‘You make quite a SIMPLE dough then.’

(17) a. [des!] **kaaesch du** selber endscheida, wie-de-s
   that can-2SG you self decide how-you.2SG-it willsch.
   want-2SG
   ‘You decide that for yourself, how you want it.’

   b. un des **kaaesch Ø** ja ed [nooçh!]kaufa.
   and it can-2SG _ well not after-buy
   ‘Well, and you can’t replace it by buying another one.’

However, there is one verb that is very common with 2SG null subjects but uncommon with overt 2SG subjects, namely *wissa* ‘to know’. The inflected form of this verb is *woisch* ‘know-2SG’ or some variant of it (*weisch, wasch, moisch, wois*). Some null examples are given in (18), (20), (21) and (22), some overt examples are given in (19) and (23).

(18) **woisch Ø**, was doo [dren!] isch?
   know-2SG _ what there inside is
   ‘Do (you) know what’s in there?’

(19) **wois-d** was der dann zo mir gsait hot?
   know.2SG-you.CL what he then to me said has
   ‘Do you know what he said to me then?’
also *woisch* Ø, i mecht [błoß!] ned, daß, wenn wenn [i!] so know-2SG _ I would just not that if if I des jetzt [ei!]schmeiß ...
it now in-throw ‘Well (you) know, I just wouldn’t like that if I post it now …’

*woisch* Ø, i han die jo [kschpritzt!] know-2SG _ I have them yes piped ‘(You) know, I PIPED them.’ (about biscuit making)

s-isch jetz grad so-s richdiche [wedder!], *woisch* Ø, it-is now just so-the right weather know-2SG _ ‘It’s just the right weather right now, (you) know.’

do han i messa a fronleichamsprozession aufstella, then have I must a corpus-christi-procession organise *wois-t*, de ganz prozession. know.2SG-you.CL the whole procession ‘Then I had to organise a Corpus Christi procession, you know, the whole procession.’

*woisch* ‘know-2SG’ is polyfunctional. It functions as a matrix clause that takes a sentential complement in utterances where the speaker puts an information question or rhetorical question to the 2SG addressee: Do you know X? Such matrix-clause *woisch* is utterance-initial, see the example in (18), with a null subject, and in (19), with an overt subject. If *woisch* were omitted here, the semantic content of the utterance would change – they would not be questions any more.

However, matrix-clause *woisch* (with or without an overt subject) is quite rare in the corpus. The six informants mostly use *woisch* as a (near-) formulaic expression, parenthetically tacked onto an utterance. Some such examples are given in (21)–(23): *You know, I piped them; Just the right weather, you know; You know, the whole procession.* If such near-formulaic *woisch* were omitted, the semantic content and the grammaticality of the utterance would not be affected. *woisch* in these cases simply functions as a discourse marker serving addressee involvement and speaker-hearer contact (cf. English *y’know, y’see*).12 Such near-formulaic *woisch* can be either clause-initial or clause-medial, often to specify

---

12 Crosslinguistically, it is not uncommon for short matrix clauses to also function as (and over time turn into) discourse markers (cf. e.g. Aijmer 2002; Traugott & Dasher 2002; Lindström & Wide 2005; Fischer 2006).
something further, or tagged on at the very end, similarly to a question tag. Some more examples are provided in (24)–(27).13,14

(24) i hab ksagt, also woisch Ø, [ihr!] sin ja aber au
I have said so know-2SG _ you.PL are well but also
[arg!] [schbääd!] [dran!]
bad _ late _ there
‘I said, (you) know you’ve really left it awfully late.’

(25) a [bekannde!] von mir, die hod, woisch Ø # äh,
an acquaintance of me she has know-2SG _ ah
zöliakie.
coeliac-disease
‘An acquaintance of mine, she’s got, (you) know, coeliac disease.’

(26) dia hod mi verschdanda, wasch Ø, dia hod alles tipptopp
she has me understood know-2SG _ she has all tiptop
gmacht.
made
‘She understood me, (you) know, she did everything tiptop.’

(27) des [betäubt!] dann au, woisch Ø, daß des ed so [weh!]
it numbs then also know-2SG _ that it not so pain
tut odr so.
does or so
‘It numbs as well, (you) know, so that it doesn’t hurt so much.’

Discourse-marker woisch is frequent in the speech of all six informants (106 out of 108 2SG woisch occurrences) and in all cases but one the informants use such woisch without an overt subject (105/106). Only one informant, incidentally the oldest one, produces woisch with an overt subject, but even she does so very rarely. Thus, the high frequency of near-formulaic woisch is the main reason for the fact that the lexical verb wissa ‘to know’ predominantly occurs with null 2SG in the corpus. However, I would like to stress that the occurrence of near-formulaic null-subject woisch cannot explain the high rates of 2SG null subject in the Swabian data in general, since the majority of 2SG null subjects do not occur with woisch but with other verbs (or with complementisers). There are 61% clause-internal null 2SG subjects overall, and if one were to subtract the near-formulaic woisch cases, there would still be 46% 2SG null subjects. The high occurrence of 2SG null subjects in the corpus

13 I use the term near-formulaic expression here because woisch is not exclusively formulaic. The speakers also use woisch non-formulaically: as a matrix clause verb taking a sentential complement.

14 The fact that the verb ‘know’ emerges as a form exhibiting unusual behavior means that Swabian presents yet another language in which know shows out-of-the-ordinary behaviour (cf. also I know not, je sais pas, etc.). Thanks to a reviewer for pointing this out.
across a wide variety of syntactic, phonological and lexical environments clearly bolsters the claim previously made in the literature that 2SG null subjects are a regular phenomenon in Swabian, except that earlier studies were not backed up by any quantitative data (e.g. Cooper & Engdahl 1989:43; Haag-Merz 1996; Fuß 2005:219; Axel 2007:321).

6. Results: First person singular null subjects

In the aggregated data from all six informants, there are 1037 contexts for a first person singular (1SG) subject to occur, including 995 clause-internal ones. Only in a small minority of these contexts is the subject null (1.3% (13/995), recall Table 4). The proportion of null 1SG subjects ranges between 1% and 3% for the six individual speakers and is thus quite homogeneous, without any discernible age-related differences. Compared to 2SG, the number of contexts for 1SG subjects is high, but null 1SG subjects are exceedingly rare. As the literature on modern Germanic dialects makes a number of claims concerning the occurrence of 1SG null subjects, it is worth investigating these from a corpus point of view.

6.1. Syntactic and lexical contexts for 1SG null subjects

Clause-internal 1SG null subjects occur postverbally in main clauses and post-complementiser in embedded clauses. Null forms are illustrated by the examples in (28)–(31).

(28) [About the speaker saving her mother in a fire]  
i be grannd, weaga der muaddr, daß Ø se no  
I am run because.of the mother that _ her still  
d’schdääga ra breng.  
the-stairs down get.1SG  
‘I ran because of mother, so that (I)’d still get her down the stairs.’

(29) [Recalling what sort of fruit trees a female acquaintance had]  
schbitzbira, des hot-se au gheed, erener Ø me dra,  
pointed-pears that has-she also had remember _ me on-it  
des hot-se au ghed.  
that has-she also had  
‘Pointed pears, those she had as well, I remember that, those she had as well.’

(30) [Leafing through a picture album, searching for a particular photograph]  
ward mal # hap-Ø-s-n no drin +…  
wait just have.1SG-_it-then still in-it  
‘Wait. Do I still have it in here then?’
(31) [Discussing whether a misspelling of a name has been corrected in the new edition of the phone book]

and now in the new have-3PL-they-it though think.1SG _
gändert

‘And now in the NEW one I think they’ve changed it though.’

All six informants produce at least one instance of a clause-internal null 1SG subject. These null subjects appear to be in free variation with unstressed overt 1SG subjects (i ‘I’), as the same speaker produces both null and overt subjects in near-minimal pairs with an identical verb form or complementiser, as illustrated in (32)-(35).

(32) a. wo Ø dr neidoa han
   that _ you.DAT in-put have-1SG
   ‘that I put in for you’

   b. wo i dr neidoa han
   that I you.DAT in-put have-1SG
   ‘that I put in for you’

(33) a. daß Ø se no d’schdääga raa breng
   that _ her still the-stairs down get. 1SG
   ‘that (I)’d still get her down the stairs’

   b. daß i des so ka
   that I it so can. 1SG
   ‘that I can do it so well’

   c. dia warad scho no sea, daß i fehl.
   they will-3PL yet still see that I miss.1SG
   ‘Come time they’ll realise that I’m missing.’

(34) a. ward mal # hap-Ø-s-n no drin +...
   wait just have.1SG-_it-then still in-it
   ‘Wait. Do I still have it in here then?’

   b. un dann hab-i-mi widdr dumpf [erinnert!]
   and then have.1SG-I-me again dimly remembered
   ‘and then I dimly REMEMBERED again’

   c. den hab-i no dr[heim!]
   that have.1SG-I still at-home
   ‘I’ve got that at HOME still.’

(35) a. un jetz in dem [neua!] hen-se-s au glaub
   and now in the new have-3PL-they-it also think.1SG
   Ø gändert
   _ changed
   ‘And now in the NEW one I think they’ve changed it though.’
There is one lexical verb that stands out as far as null 1SG subjects are concerned, glauba ‘to think/believe’. Half of the clause-internal null 1SG subjects (7/13) occur with the inflected form glaub think.1SG ‘I think’. Interestingly, the use of glaub with null subject is near-formulaic and reminiscent of the discourse-marker use of woisch ‘know-2SG’ with null 2SG subject (recall section 5.3). Like woisch, glaub is polyfunctional. It can serve as a matrix clause taking a sentential complement (I believe that X), but in such contexts the informants use overt subjects. They do not employ glaub + Ø as a matrix clause but rather like a discourse marker, a parenthetical I think comment and hedge on the proposition of a clause, as in (35), where the speaker hedges her propositions: In the new phone book, a faulty entry has been changed, but this is not an established fact but only what the speaker thinks (35a); A friend once brought some cake to have with the coffee, at least that is what the speaker thinks (35b). Instead of discourse marker glaub + Ø, one can also use glaub + i, i.e. the same verb with an overt postverbal 1SG subject, as in (35c), but this only happens once in the corpus. A majority of the clause-internal 1SG null subjects thus occur in the context of a semi-formulaic expression. It would be interesting to know whether postverbal null subjects in semi-formulaic expressions also occur in other non-standard varieties of Germanic; previous studies of Swabian and related German dialects have not mentioned such occurrences as far as I am aware.

6.2. Phonological conditions on 1SG null subjects?

Earlier non-corpus-based studies of South German dialects have instead discussed the phonological context of 1SG subjects. It has been suggested that the possibility of dropping 1SG subjects might be dependent on the phonological form of the subsequent element. Clitic clusters have featured prominently in this discussion, since they appear to create a context from which 1SG subjects can be dropped in Zurich German, Bernese German and Swabian (e.g. Cooper & Engdahl 1989:38–39; Nübling 1992; Haag-Merz 1996; Werner 1999). Clitic cluster may be something of a misnomer here: The researchers refer to a context in which an unstressed overt 1SG subject (‘clitic’) could have occurred postverbally or post-complementiser, followed by at least one other
unstressed element, such as an object pronoun (‘clitic’). They propose that in such a context, 1SG subjects are susceptible to deletion in the phonological component. The main idea here is that when a strong syllable is followed by two weak syllables, the weak syllable in the middle has a tendency to phonologically reduce to the point of disappearing completely. In the present context, the strong syllable would be a monosyllabic or multisyllabic end-stress 1SG verb, the first weak syllable a 1SG pronominal subject, and the subsequent weak syllable another unstressed element (cf. Nübling 1992:270–271; Werner 1999:128).

For concreteness, let’s look at the example in (34b) hab-i-mi ‘have-I-me’, which has the constellation strong-weak-weak (hab-i-mi). In such a context, the weak syllable in the middle should be prone to deletion, leading to the appearance of 1SG subject drop (hab-Ø-mi). Haag-Merz (1996) discusses Swabian 1SG subject drop from this angle, but with a little twist, based on her own native acceptability judgments. According to her, there exists only one context in Swabian where 1SG subjects can be null, namely in medial position in clitic clusters, particularly those that contain an object clitic (1996:161–166). Moreover, she claims that a 1SG subject must be null if it is followed by a vowel-initial clitic, to avoid a vowel-vowel hiatus between the 1SG subject consisting of a vowel (-i/-e ‘I’) and an element that also begins with a vowel or a syllabic sonorant, such as the object clitic (e)n ‘him.ACC’ or (e)m ‘him.DAT’ (Haag-Merz 1996:162–163). The occurrence of 1SG null subjects would thus be phonologically determined. Native speakers of Swabian (like myself) may disagree with Haag-Merz’s judgments, but here I simply want to look at the corpus data in light of her claims.

As will be recalled, very few 1SG subjects are null (1.3%, 13/995), and more than half of these occur with semi-formulaic glaub (7/13), which is not a context mentioned in Haag-Merz (1996). 4 of the 6 remaining 1SG null subjects in the corpus occur in clitic clusters. We have already seen these examples in (28)–(30) and (32): daß Ø se, erener Ø me, hap-Ø-s-n, wo Ø dr. In none of these cases would an overt 1SG pronoun have given rise to a hiatus between two vowels (daß-e-se, erener-e-me, hab-e-s-n, 

\[15 \text{In Swiss German dialects, 1SG null subjects may also occur in the context of an article that is homophone with an object pronoun (e.g. Cooper & Engdahl 1989:39; Werner 1999:50, 128 for Zurich German).}

\[16 \text{According to Haag-Merz, examples (i)-(ii) with an overt 1SG subject are ungrammatical. The 1SG subject must be null (Haag-Merz 1996:161-162, her examples (215b), (219b)).}

(i) *geschtern han-e-n verschompfe.
   yesterday have.1SG-I-him.ACC told-off
   ‘Yesterday I told him off.’

(ii) *geschtern han-e-m a bifle gholfe.
    yesterday have.1SG-I-him.DAT a bit helped
    ‘Yesterday I helped him a bit.’
null subjects in Swabian), which according to Haag-Merz is the phonological reason for the subject to be dropped.

In order to compare these 4 instances of null 1SG subjects in clitic clusters with overt 1SG subjects, I investigated all clitic clusters in the corpus. There were 134 clitic clusters with an overt 1SG subject (cf. e.g. (32b), (34b)). In 18 of these, a vowel-vowel hiatus did occur due an overt 1SG subject (i or e ‘I’) and a subsequent vowel-initial element, and could have been avoided if the subject had been dropped. However, this is not what the informants chose to do; they did pronounce the 1SG subject. This finding suggests that Haag-Merz’s condition that 1SG subjects must be dropped for phonological reasons to avoid vowel-vowel hiatus does not hold for the six informants here. (It is of course still possible that her proposed phonological condition could hold for other speakers, such as Haag-Merz herself.)

In general, the informants drop very few 1SG clause-internal subjects. Since this is a corpus study and not an introspective one, we do not know what these informants could and would allow in an acceptability judgment task. I must leave this issue for future research. What we do know however is that the informants exhibit an extremely strong preference for 1SG subjects to be overt (98.7%) in spontaneous production, and this in itself is a noteworthy finding. Earlier (non-quantified) studies of Swabian and related non-standard varieties of German have focused on the optionality between null and overt 1SG subjects and presented null 1SG subjects as a phenomenon that is commonplace. This may be slightly misleading: in spontaneous production, 1SG subjects in the Swabian corpus data are virtually always overt. It would therefore be useful if future studies of other varieties were to include quantitative data as well. It might also be interesting to investigate 1SG null subjects in Swabian from a diachronic perspective, as it is possible that 1SG null subjects (e.g. in clitic clusters) were more frequent in the past and that the present corpus only shows the remaining vestiges of this. Such speculations could be tested if older audio-recordings of Swabian speakers were to be compared with the recordings of the informants here.

7. Results: Third person singular neuter subjects

In the aggregated data from all six informants, there are 1084 contexts for a third person singular neuter (3SGNEU) referential subject to occur, including 1005 clause-internal ones, and only in a small minority of these contexts is the subject null (2.5% (25/1005), recall Table 4). The proportion of null 3SGNEU subjects ranges between 1% and 13% for the individual speakers, so there is a fair bit of variation, and interestingly, this appears to be age-related. The older speakers omit clause-internal 3SGNEU subjects 6%–13% of the time, whereas the three younger speakers only do so 1% of the time (see below). The number of
contexts for 3SGNEU subjects is comparatively high, on a par with those for 1SG and nearly three times as high as those for 2SG. However, when compared to the high proportion of subject drop in 2SG contexts discussed in section 5, the proportion of 3SGNEU null subjects is very low (2.5%) and on a par with that of 1SG drop (section 6). Incidentally, 3SG masculine and 3SG feminine subjects are never dropped. In what follows, the realisation of clause-internal 3SGNEU subjects will be investigated in more detail.

3SGNEU null subjects occur in both main and embedded clauses. However, all 25 instances of null 3SGNEU occur after one and the same finite verb, *isch* ‘is’, i.e. the present tense 3SG form of the copula verb *sei* ‘to be’, or after the complementiser *daß* ‘that’. This is in striking contrast to 2SG and 1SG null subjects which, as will be recalled, occur with a wide variety of verbs and complementisers. Some examples of clause-internal 3SGNEU drop are given in (36)–(38).

(36) [About putting crystallised sugar in her drink]
nee, i brobier-s erschmool mid [oim!] +... *daß* Ø dann ned
no I try-it first with one that _ then not
zu [süß!] wirt.
too sweet gets
‘No, I’ll try it with ONE (lump) first, so that (it) doesn’t get too sweet.’

(37) [About an ache in her foot]
Wenn-e hald ned [drauf!]kschdanda ben, doo *isch* Ø a
when-I well not on-it-stood am then is.3SG _ a
weile [besser!]
while better
‘When I haven’t stood on it, then (it)’s BETTER for a little while.’

(38) [About an ache in her foot]
Und [obens!] *isch* Ø hald dann gell viel [schdärker!]
and evening is.3SG _ well then well much stronger
‘And in the evenings, well then (it)’s a lot stronger.’

Even though 3SGNEU null subjects are exclusively found after *isch* and *daß*, the null form is not obligatory here, as the speakers also produce overt subjects in the very same contexts, as shown by the near-minimal pair examples in (39)–(40) from the same speaker.17

17 One reviewer wonders whether these instances of clause-internal 3SGNEU null subjects shouldn’t be treated as null expletives. I would say no, as these subjects are referential, referring for example to a hot drink (36), a foot ache (37)–(38), or a meadow (40). The same reviewer asks whether the Swabian 3SGNEU null subjects would be grammatical in Standard German. The answer is no; all 25 cases would require an overt subject in Standard German.
Why then might 3SGNEU null subjects be restricted to post-isch and post-daß?

My answer will be a phonological one. Both isch /ɪʃ/ and daß /das/ end in a sibilant. Overt 3SGNEU clitic subjects also consist of a sibilant, syllabic es /əs/ or non-syllabic s /s/. When such a subject follows isch or daß, this results in strings like isch-es /ɪʃəs/, isch-s /ɪʃs/, daß-es /dasəs/, daß-s /dəs/. It is quite conceivable that for ease of pronunciation one might avoid sibilant clusters like /ss/ and /ss/. This can easily be done, either by breaking up the cluster via resyllabification into /ɪʃ/ + /əs/ and /das/ + /əs/ respectively, particularly so as Swabian has es /əs/ as a 3SGNEU form, or by deleting the second sibilant (/s/) in the cluster. The (e)s may underlingly be there but be phonologically deleted at post-syntactic level. Such deletion will result in a null 3SGNEU subject.18

I investigated this possibility for the six speakers in the corpus. In total, they produced the aforesaid 25 cases of isch/daß +Ø but also 30 cases of isch/daß + overt 3SGNEU. They also produced hundreds of overt 3SGNEU

18 One reviewer asks whether we would not expect other varieties of German potentially to exhibit similar 3SGNEU –s deletion if phonology is at issue here. I think we would indeed expect this, and corpus studies of other spoken varieties would therefore be most welcome. Note however that phonologically mediated deletion is just one potential option here, another is resyllabification (see above). Moreover, phonological differences between Swabian and other varieties will play a role: Depending on which morpho-phonological form the preceding element and the 3SGNEU subject takes, sibilant clusters might not even arise and hence not be prone to deletion. For instance, the 3SG copula plus 3SGNEU subject in Swabian is isch-s /ɪʃs/ and realisation of the sibilant cluster /ʃs/ requires a change of place of articulation from alveolar-palatal to alveolar. By contrast, the corresponding Standard German form is ist es /ɪstəs/, with no consonant cluster at all, or ist-s /ɪstʃs/, where the cluster /ts/ is much easier to pronounce than /ʃs/, as /ʃs/ requires no change of place of articulation and moreover corresponds to a high-frequency German affricate. It is likely then that in a variety of German similar to Standard German, the 3SGNEU subject in ist-s would not be deleted.
subjects in other contexts, but no null forms outside the context of *isch* and *daβ*. Interestingly, the speakers exhibited divergent individual preferences: Three speakers nearly always omitted the subject and thus avoided a sibilant cluster arising, whereas the other three nearly always produced an overt clitic subject that resulted in a sibilant cluster. In a few cases, they produced an overt subject and avoided the sibilant cluster by resyllabification. As shown in Table 5, the older informants omitted the subject 91% in the context of *isch/daβ*, whereas the younger informants produced 87.5% overt subjects in the same context.

Both young and old informants also produced many instances of the full form of the 3SGNEU subject, *des ‘it’* in post-*isch* and post-*daβ* contexts. These are not included in the counts. What separates the informants from each other is their preferential use of the null form versus the (*e*)s clitic form of the 3SGNEU subject (Table 5).

It is possible that there is a language change going on in Swabian; however, I do not want to make too much of the age divide, since we are looking at six informants only and the pattern observed may simply be due to individual preference and variation. Further studies could resolve this issue, particularly if older audio-recordings are included (such as the extensive recordings made under the auspices of Arno Ruoff in Tübingen in the 1950s and 1960s, now at the Deutsches Spracharchiv in Mannheim).

What we can say is that the omission of 3SGNEU subjects is qualitatively very different from the omission of 2SG and 1SG subjects discussed in sections 5 and 6.

### 8. Discussion and conclusion

This paper has explored the distribution of 4,000 null and overt subjects in finite clauses in a new spontaneous production corpus of Swabian. Few studies have investigated referential null subjects in modern non-standard German vernaculars, and to my knowledge none has done so on the basis of quantified data, so the present study of Swabian is a step towards filling this empirical gap.

Averaged across all persons and numbers, there are 6.8% clause-internal referential null subjects in the corpus, but this figure does not tell us much. Rather, the corpus data indicate that the null subject

### Table 5. 3SGNEU subjects after *isch* and *daβ*

<table>
<thead>
<tr>
<th></th>
<th><em>isch/daβ + Ø</em></th>
<th><em>isch/daβ + es/s</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>Elderly informants</td>
<td>91% (21/30)</td>
<td>9% (2/30)</td>
</tr>
<tr>
<td>Younger informants</td>
<td>12.5% (4/32)</td>
<td>87.5% (28/32)</td>
</tr>
</tbody>
</table>
phenomenon in Swabian is highly selective and only occurs for part of the verbal paradigm: 2nd person singular (2SG), 1st person singular (1SG), and 3rd person singular neuter (3SGNEU). These results largely confirm the findings of an older, but unquantified study, which describes 2SG and 1SG subject drop (though not 3SGNEU) in Swabian (Haag-Merz 1996). A closer investigation of the three types of subject person reference in the present corpus reveals clear differences in frequencies and distributional patterns: Null 2SG subjects are found in the majority of 2SG contexts (61%), whereas null forms are rare in contexts for 1SG and 3SGNEU (1.3% and 2.5%, respectively).

Swabian null 2SG subjects appear to be in free variation with overt pronouns (section 5). A functional analysis of 2SG subjects in the corpus indicated that whilst focus/contrast on the subject is expressed via a stressed pronoun, old information/continuous topic is expressed either via an unstressed overt or a null 2SG subject. No difference in discourse-pragmatic function could be discerned between the unstressed overt and the null 2SG form.

Null 2SG subjects are regularly produced by all speakers and in a large variety of syntactic, phonological and lexical contexts: postverbally in main clauses (frequently) and post-complementiser in embedded clauses (less frequently), and with many different types of verbs. One verb, woiisch 'know-2SG', is especially frequent with a null 2SG subject, and this happens specifically in contexts where the verb is not used as a matrix clause but as a semi-formulaic discourse marker. However, as such instances with woiisch only make up a fraction of null 2SG and as the majority of null subjects in the corpus occurs in a wide range of other contexts, I have argued that 2SG subject drop is a robust phenomenon in Swabian which cannot be dispensed with by relegating it to semi-formulaic expressions. I will leave it open how the null 2SG subject property in Swabian should be modelled theoretically.

Turning to null 1SG subjects, these are also used by all speakers, but much less frequently than null 2SG (section 6). Half of the null 1SG subjects occur in a particular lexical context, namely together with glaub 'think.1SG', where the verb functions as a semi-formulaic hedging discourse marker. The remaining null 1SG subjects occur post-verbally and post-complementiser, but here it has proved hard to find any common denominator for them. I compared the Swabian null 1SG data to earlier proposals in the literature, according to which 1SG subject drop is phonologically conditioned and restricted to medial position in certain clitic clusters (Nübling 1992; Werner 1999; and in particular Haag-Merz 1996). The phonological constraints proposed there could however not be supported by the corpus data. In general, 1SG subject drop in present-day Swabian appears to be a marginal phenomenon, as only 1% of the clause-internal 1SG subjects are dropped. This is much rarer than has previously been suggested.
Turning to 3SGNEU referential null subjects, these are also attested for all speakers, but again much less frequently than null 2SG. Crucially however, all of the clause-internal null 3SGNEU subjects occur in a particular lexical context, namely after the finite verb *isch* ‘is’ or after the complementiser *daß* ‘that’. What these two elements have in common is that they end in a sibilant, and thus null 3SGNEU subjects are restricted to a particular phonological context. The occurrence of 3SGNEU null forms in Swabian has not been noted in the literature before. I compared the null forms to overt 3SGNEU subjects in the context of *isch* and *daß* and suggested that the null form arises as a consequence of post-syntactic phonological simplification processes (section 7).

The findings from Swabian speak to a number of proposals in the literature on null subjects in general. They support the view (nowadays voiced by many) that the earlier binary division of languages into pro-drop versus non-pro-drop is too simplistic; rather, there are also quite a number of languages that allow partial pro-drop, i.e. referential null subjects for certain person-number constellations, tenses, etc., and only for those (e.g. Huang 1995; Vainikka & Levy 1999; Cole 2009; Rosenkvist 2009). Regarding these constellations, I would expect cross-linguistic variation concerning which person/number/gender etc. allows null subjects, and the situation found for Swabian is just one such constellation. I am therefore sceptical of attempts to unify partial null-subject by proposing some narrow defining characteristic of partial null-subject languages, such as generic pronouns having to be null (e.g. Holmberg, Nayudu & Sheehan 2009; Roberts & Holmberg 2010:10–12).19

The Swabian data also add to our knowledge of null subjects in Germanic, where it has long been claimed that clause-internal referential null subjects are ungrammatical, but where recent dialectology and diachronic research has shown that many Germanic varieties do allow null subjects, again for certain person-number constellations or in specific morpho-syntactic contexts only (e.g. Werner 1999 on Zurich German; Axel 2007 on Old High German; Volodina 2009 on Early New High German).

The findings from Swabian also speak to more specific proposals regarding null subjects in non-standard Germanic V2 languages. One is the claim that complementiser agreement might be the licensing force behind referential null subjects, such that only those varieties that have inflected complementisers (e.g. Bavarian and Frisian) allow subjects to be null in embedded clauses, i.e. in post-complementiser position (Bayer 1983/1984; Hoekstra 1997; Weiß 1998, 2005; Fuß 2004; Axel 2007; Axel & Weiß 2010). The results from Swabian do not lend support to this, since Swabian lacks complementiser agreement but null subjects in embedded clauses are attested all the same, though not at the high rate as

19 Recall that generic subject pronouns are never null in the Swabian corpus (Table 4).
in main clauses as far as 2SG subjects are concerned. At present, I do not have a cogent explanation for the different rates of null 2SG subjects in Swabian main clauses (71% null) versus embedded clauses (14% null, see section 5.1). One should perhaps investigate diachronic continuity here, i.e. the possibility that present-day Swabian shows remnants of an older pattern where referential null subjects were restricted or largely restricted to main clauses. Eggenberger (1961) and Axel (2007: ch. 6) have shown for Early Old High German texts that referential subjects were dropped near-exclusively in main clauses and hardly ever in embedded clauses, and Volodina (2009:59–62) shows the same for autobiographic Early New High German texts. However, these authors studied written genres, which may not be comparable to informal oral Swabian. Moreover, Eggenberger’s, Axel’s and Volodina’s sources do not contain enough contexts for 2SG subjects, so we do not know whether there indeed was an asymmetry between main and embedded clauses for 2SG null subjects in older versions of German, or whether the asymmetry mainly concerned 1st and 3rd person subjects where more data exists (Axel 2007:315; Volodina 2009:58–59).20,21

I have also looked into the possibility that the realisation of null versus overt subjects in Swabian might be pragmatically conditioned.22 Here, a number of researchers have suggested that null subjects encode old information and/or continuing topic, whilst overt subject pronouns signal new and/or focused information or change of topic (e.g. Carminati 2002; Öztürk 2002; Frascarelli 2007; Cole 2009). For modern Germanic vernaculars in particular, Rosenkvist (2009) has argued that null subjects are used as a default whilst overt pronominal subjects trigger emphatic or contrastive interpretations. However, in the Swabian corpus data, there are no discernible semantic or pragmatic differences in the interpretation of null vs. unstressed overt 2SG subjects. Focus/contrast on the 2SG subject is expressed by using a stressed 2SG pronoun and not by turning a ‘default’ null form into an (unstressed) overt subject pronoun. Old information/continuous topic is either expressed via an unstressed overt or a null 2SG subject (section 5.2). No difference in discourse-pragmatic

20 Held (1903: 89–92, 129–138) provides numerous examples of 2SG subject drop from late Middle High German and Early New High German (ENHG) sources and states that 2SG subjects in ENHG were “frequently” dropped not only in main clauses, but also from within relative clauses and $d$-clauses. However, Held provides no actual figures that would allow us to investigate the potential asymmetry between main vs. embedded clauses in ENHG.

21 Incidentally, Old Swedish exhibits a similar asymmetry between main vs. embedded clauses concerning referential null subjects, though only for 3rd person, as the figures for the other persons, including 2SG, are again very low (Håkansson 2008:101–102).

22 I have only investigated this for the case of Swabian 2SG subjects where there are many null forms. For 1SG and 3SGNEU subjects, which are nearly always overt, a pragmatic explanation of null subject can be rejected straightforward: It is hardly likely that 98.7% and 97.5% of the 1SG and 3SGNEU subjects are overt because they encode new/focused information or change of topic, and that only 1.3% or 2.5% of them encode old information or continuous topic and are therefore null.
function could be discerned between the unstressed overt and the null 2SG form. It is therefore unlikely that discourse-pragmatic factors determine the use of null vis-à-vis overt 2SG forms in Swabian. I am aware that the existence of optionality and free variation, i.e. the possibility that two constructions with the same meaning and function coexist in one language, is often disputed in syntactic theory. The empirical facts, however, appear to suggest otherwise: native-speaker grammars do allow optionality.

The data can also be related to another proposal in the literature, namely that referential null subjects are only possible if there is a verb form that can distinctly and uniquely identify the omitted subject (e.g. Hyams 1986; Rohrbacher 1994; Duarte 2000 and many others). This appears to be generally true for Swabian, as all three person/number constellations for which we have found null subjects (1SG, 2SG, 3SG) have distinct subject-agreement verb morphology, as shown by the verbal paradigms in (41).

(41) komma ‘to come’ doa ‘to do’ hau/han ‘to have’ sei ‘to be’
1SG komm dua hau/han be
2SG kommsch duasch hosch bisch
3SG kommt duat hot isch
1PL kommet dend hend sen(d)
2PL kommet dend hend sen(d)
3PL kommet dend hend sen(d)

In the 2nd person singular, the inflection on the verb is –sch (e.g. kommsch), an ending that does not occur anywhere else in the paradigm. For 1SG, the inflected verb takes the form of the stem (e.g. komm), and for 3SG, the inflection is –t (e.g. kommt). Neither desinence is identical with another desinence in the present tense paradigm in (41). One may thus want to assume a connection between distinct verb agreement and the possibility of referential null subjects. However, the 1SG forms are homophonous with the imperative form, and the 3SG inflection –t is also found on past participles, which may go against an uniquely identifying inflection approach. The 3SG copula isch ‘is’ as well as preterite-present modal verbs do not carry the typical 3SG –t inflection. Note also that there is no distinct 3SG inflection for grammatical gender; nevertheless, as we have seen, Swabian only allows 3SG neuter subjects (but not 3SG feminine and masculine subjects) to be null in specific contexts. This is unexpected under a unique-inflection approach to (partial) null subject.

Whilst a unique verb form may be necessary to license a referential null subject in Swabian, it is not sufficient; if it were, we might expect more

---

23 A similar point about free variation concerning a phenomenon other than null subject is made in the literature on periphrastic do constructions in historic and present-day West Germanic language varieties (cf. e.g. Bohnacker (in press)).
1SG and 3SG null subjects than are attested in the corpus, and we might expect them across all sorts of contexts. Yet as we have seen, the occurrence of 1SG and 3SG null subjects – in contrast to 2SG – is quite restricted, both lexically and phonologically. For ideas other than distinct verbal inflection licensing of partial null subject see e.g. Biberauer et al (2010).

As I have tried to show, the null subject phenomenon is highly selective; not only does it depend on person/number/gender information, but it may also depend on the local context of the subject. The Swabian data suggest that depending on person/number/gender, different conditions for null subjects may be operative in the same language at the same time.

References


Null subjects in Swabian


Received October 1, 2010
Accepted September 20, 2012

Ute Bohnacker
Dept of Linguistics and Philology
Uppsala University
Box 635
75126 Uppsala
Sweden
ute.bohnacker@lingfil.uu.se

© 2013 The Editorial Board of Studia Linguistica