Complement clause structure immune to phonological influences in German but not in English

Finite complement clauses (CC) in German and English come in two varieties, viz. i) those that are introduced by a complementiser (hence introduced CC) and ii) those that forego a complementiser (unintroduced CC). In German, introduced CCs display verb-final syntax, which is characteristic of subordinate clauses in German. The word order of unintroduced CCs (a.k.a dependent main clauses; Auer 1998) corresponds to the syntax of simple declarative clauses with the tensed verb in second position (V2). English, in contrast, does not show a word order difference between introduced and unintroduced CCs, apart from the presence or absence of the complementiser.

Several studies on English suggest that the choice between introduced and unintroduced CCs is, among other things, conditioned by phonology: Jaeger (2006) and Lee & Gibbons (2007) suggest that speakers tend to omit the unstressed complementiser when the word at the top of the CC starts in an unstressed syllable (avoidance of stress lapse). Walter & Jaeger (2005) show a clear effect of phonological identity avoidance (OCP): if the top of the CC has the same word form as the complementiser, the CC is preferably unintroduced (avoidance of double *that* sequences).

By means of a language production experiment and two corpus studies we show that the choice between introduced and unintroduced CC in German is immune to phonological influences of the type found in English.

Experiment 1a and 1b: Recalled language production (modelled on Lee & Gibbons 2007)

We presented participants with 32 written test sentences like (1) or (2), with 8 different embedding verbs that allow for both introduced and unintroduced CCs (*denken, finden, glauben, hoeren, hoffen, meinen, sagen, wissen* `think, find, believe, hear, hope, reckon, say, know'). Participants were asked to recall each sentence after a distractor task. In Exp 1a (32 participants) the distractor task was recalling another (unrelated) sentence; in Exp 1b (32 participants) it was solving a simple arithmetic task. The 8 conditions systematically varied in terms of the rhythmic environment at the clause boundary. Based on the results by Lee & Gibbons, we predicted more recalled CCs introduced by unstressed *dass* when the embedding verb ends in a stressed as opposed to unstressed syllable, and when the CC subject started in a stressed as opposed to unstressed syllable. Conversely, we expected more unintroduced CCs when syllables around the clause boundary were unstressed.

(1) Felix glaubt(e), dass (Nádja/Nadíne) den Brief geschrieben hat.
(2) Felix glaubt(e), (Nádja/Nadíne) hat den Brief geschrieben.
Felix thinks/thought (that) Nadja/Nadine has written the letter

Both experiments (valid data points Exp 1a: 470 (46%); Exp 1b: 973 (95%)) show a clear preference for recalls involving introduced CC but failed to produce the predicted rhythmic effects. By means of Bayes Factor analysis (see Nicenboim & Vasishth 2016, for a tutorial), we provide strong evidence (with Bayes factors >10) in favour of the null hypothesis according to which the rhythmic environment does not affect the choice between introduced and unintroduced CCs in German.

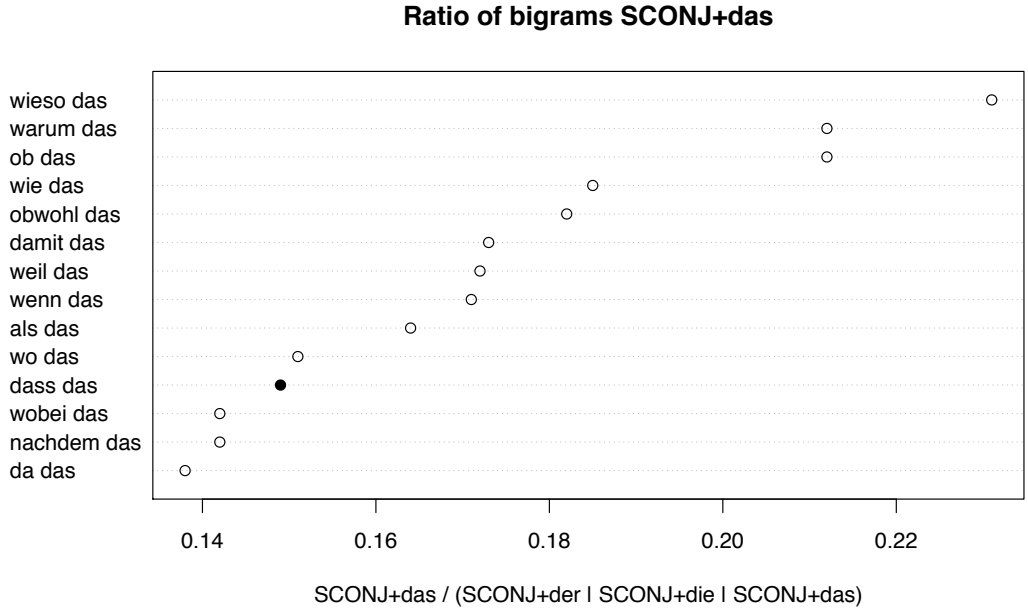
Experiment 2: Corpus study on rhythm and CC structure (TÜPP/DZ corpus)

In order to validate the findings of Exp 1, we searched the taz newspaper corpus (TÜPP/DZ) for all tokens of the 8 embedding verbs that were immediately followed by a complement clause (with or without complementiser) with a proper name as clause-initial subject. This search yielded 2751 complement clauses, 1476 subordinate clauses with, and 1275 subordinate clauses

without, the complementiser *dass*. Against predictions, but in line with Exp 1, there was no evidence for introduced CCs (with *dass*) to be more likely when the proper name started in a stressed syllable (52% introduced CCs) as opposed to an unstressed one (56% introduced CCs).

Experiment 3: Corpus study on word-form OCP and CC structure (DeReKo)

In an attempt to replicate the findings of Walter & Jaeger (2005; avoidance of double *that* sequence) for German, we searched the German Reference Corpus (DeReKo-W, written section) for bigrams involving various subjunctions followed by the definite determiners *der, die* or *das*. We predicted that if CC structure is susceptible to phonology, we should observe a similar OCP effect as Walter & Jaeger did, i.e. a clear avoidance of *dass das*. The plot in the Figure below depicts the frequency ratio of subjunction+*das/*subjunction+(*der*|*die*|*das*) for 14 subjunctions (POS-tag "SCONJ"). While the *dass das* bigram (black dot) is at the lower end of the spectrum (.15), it is clearly not exceptionally low. Note also that the frequency ratio of singleton *das/(der*|*die*|*das)* is even lower (.14).



In sum, the apparent lack of phonological influences suggests that CC structure (introduced vs unintroduced) in German is fixed by syntax before phonology gets access to it. In contrast, the presence/absence of the overt complementiser in English is clearly affected by the phonological environment. Its absence may be conceived as phonological ellipsis. Hence, in spite of the differences concerning phonological influences on overt CC structure, the syntax proper remains phonology-free in both languages.

References

Auer, Peter. 1998. Zwischen Parataxe und Hypotaxe: abhängige Hauptsätze im gesprochenen und geschriebenen Deutsch. *Zeitschrift für germanistische Linguistik* 26(3). 284–307.

Jaeger, T. Florian. 2006. Phonological optimization and syntactic variation: The case of optional that. In *Proceedings of the annual meeting of the Berkeley Linguistics Society*, vol. 32. 175–187.

Lee, Ming-Wei & Julie Gibbons. 2007. Rhythmic alternation and the optional complementiser in English: New evidence of phonological influence on grammatical encoding. *Cognition* 105(2). 446–456.

Nicenboim, Bruno & Shravan Vasishth. 2016. Statistical methods for linguistic research: Foundational ideas (part II). *Language and Linguistics Compass* 10(11). 591–613.

Walter, Mary & T. Florian Jaeger. 2005. Constraints on optional that: A strong word form OCP effect. *Proceedings from the annual meeting of the Chicago Linguistic Society*. Vol. 41. No. 1.