On the need for spoken corpora in grammatical description:
the case of Polish *że* as an elaboration marker

Introduction The paper examines Polish *że* 'that' in its under-researched non-canonical use as a marker of elaboration. We argue that two homophonous elements need to be distinguished: the canonical ***że1* – the declarative 'that'-type complementizer** operating at the syntactic level, and ***że2* – the elaboration marker** operating at the discourse level (therefore, a discourse marker/connective). We provide evidence for the above claim showing that *że1* and *że2* behave differently with respect to a number of syntactic or discursive facts.

Phenomenon and methodology *Że2* is illustrated in (1) and (2). *Że2*-clauses elaborate on the preceding discourse by providing more detail (1), justifying, explaining, exemplifying, or simply reformulating (2). Since this usage is limited to colloquial speech, our data come from the spoken corpus Spokes (spokes.clarin-pl.eu). We have so far inspected manually circa 13,000 occurrences of *że* and retrieved 106 tokens of *że2* (work in progress).

(1) (providing more detail)
    *a w ogóle będę pracować znowu nie? **że** roznosić te takie ulotki*
    'And, by the way, I'll be working again, right? **Like/that is/namely**, distributing those leaflets'

(2) (reformulation)
    *ogólnie Internet wtedy zwalnia strasznie **że** strasznie zamula nie?*
    'In general, the Internet then slows down terribly. **Like**, really gets bogged down, right?'

Syntactic and discourse-related differences In contrast to *że1/że1*-clauses, (i) *że2*-clauses are not selected by matrix expressions, instead they are loosely appended to elaborate on the preceding discourse; (ii) *że2* is compatible with predicates that select *czy* 'if' (*pytać* 'ask') or *żeby* (*zachęcać* 'encourage') without entering a complementation relation; (iii) *że2* is an optional element; (iv) *że2*-clauses cannot be fronted; *że2* cannot host auxiliary person-and-number clitics (e.g. *że+m/ś*); (v) *że2* is more flexible in that it introduces material other than finite clauses, e.g. an infinitival clause in (1) and a noun phrase in (3); (vi) *że2* extends to non-assertive speech acts such as direct questions or imperatives, as in (4), thus it is a discourse element modifying the speech act level.

(3) (noun phrase)
    *Okazja która będzie satysfakcjonująca nie tylko na poziomie finansowym ale też rozwojowym **że** na przykład praca akademicka czy coś takiego*
    'An opportunity that will be satisfactory not only on the financial but also developmental level ***że2*** for example academic work or something like that'

(4) (question)
    *ktoś tam się pyta "No jak tam?" **że** "Co tam u was?" nie?*
    'Someone asks "Well, how are things?" ***że2*** "How are you?", right?'

Conclusion Polish *że* represents (at least) two different elements: a complementizer and an elaboration marker (see Germanos 2013 for a parallel dualism of a complementizer form in Lebanese Arabic). *Że2* does not seem to be an innovation, as it is recorded in 16th-century texts and entered in *Słownik polszczyzny XVI wieku*. It has survived only in casual speech and has

been banned from standard Polish, which opts for other (lexical) means to express elaboration. As such, the case of *że* underlines the need for the analysis of spoken data in grammatical description. In the paper, we will also mention briefly two other relic functions of *że* present in contemporary spoken Polish.

References

Germanos, Marie Aimée (2013) From complementizer to discourse marker: The functions of *ʔənno* in Lebanese Arabic. In J. Owens &A. Elgibali (eds.) *Information Structure in Spoken Arabic*, London and New York: Routledge, 145–164.

Słownik polszczyzny XVI wieku, Maria Renata Mayenowa, Franciszek Pepłowski (eds.) (vol. 1–34), Krzysztof Mrowcewicz, Patrycja Potoniec (since vol. 35 to entry ROWNY), Wrocław: Ossolineum, (1966–1994), Warszawa: IBL PAN (1995–).